(72) Inventors:
• Saito, Yasushi
Chiba-shi, Chiba (JP)
• Iwasaki, Akio
Tsuchiura-shi, Ibaraki (JP)
• Arai, Koichi
Urawa-shi, Saitama (JP)
• Yamazaki, Hiroyuki
Higashimurayama-shi, Tokyo (JP)

(74) Representative: Wächtershäuser, Günter, Prof. Dr.
Patentanwalt,
Tal 29
80331 München (DE)

Remarks:
The applicant has subsequently filed a sequence listing and declared, that it includes no new matter.

(54) **Novel LDL receptor analog protein and the gene coding therefor**

(57)   The present invention is drawn to the gene of a novel LDL receptor family receptor which participates in lipoprotein metabolism, a critical factor that triggers the onset of arteriosclerosis.

The invention provides DNA having a nucleotide sequence as shown by Sequence ID No. 1 or No.5 is disclosed as well as rabbit tissue or human tissue LDL receptor analog protein having an amino acid sequence of Sequence ID No. 2 or 6 coded by such DNA.

EP 0 773 290 A2

Description

Background of the Invention:

5   1) Field of the Invention

The present invention relates to a novel LDL receptor analog protein having a structure similar to that of LDL receptors that are responsible for the homeostasis mechanism of intracellular cholesterol and extensively participates in serum lipid metabolism, which is a critical factor that triggers the onset of arteriosclerosis. The invention also relates to
10   the gene coding for the protein.

2) Description of the Related Art

Abnormality in serum lipid metabolism is one of the most critical risk factors in the onset and progress of arterio-
15   sclerosis. Serum lipids, together with apolipoproteins, are transformed into lipoproteins primarily in the liver, secreted therefrom, transported by blood, and taken up by a variety of tissue cells.
Uptake of lipoproteins into cells occurs primarily by the mediation of receptors of respective lipoproteins. It is known that low density lipoproteins (LDL), which are taken into cells by specific membrane receptors, called LDL receptors, are metabolized within the cells and utilized as cell membrane components or similar substances. Detailed analysis of
20   familial hyperchlolesterolemia, which is a genetic disease accompanied by notable hyperchlolesterolemia due to abnormality of LDL receptors, has clarified details of the mechanism of homeostasis achieved by LDL receptors with respect to intracellular cholesterol.
It has been suggested that living bodies have not only LDL receptors but also cell membrane receptors that recognize other lipoproteins. From analyses of WHHL rabbits, which are model animals lacking LDL receptors, it was found
25   that receptors which takes principally apo-E-containing lipoproteins as ligands (remnant receptors) are present in the liver. It is also predicted that there may be HDL receptors whose ligands are high density lipoprotein (HDL). However, to date, details of the structures and functions of these receptors have not yet been elucidated. It has also been known that foaming of macrophages plays an active role in the formation of atherosclerosis, is deeply participated. Macrophages foam by taking up modified LDL —not normal LDL— which have undergone oxidation, acetylation, or glycation.
30   There have recently been discovered receptors to modified LDL which are called scavenger receptors. The scavenger receptors have been identified to be membrane receptors that have a structure completely different from that of LDL receptors.
Recent research using molecular biological techniques has identified the genes of LRP (LDL receptor-associated protein), gp 330, and VLDL receptors. The receptors have been found to have structures very similar to those of LDL
35   receptors. From analyses of these receptors, it is believed that a plurality of lipoprotein receptors are present in living bodies, and that they are closely related to lipid metabolism. LDL receptors studied in detail by Brown and Goldstein [Brown, M.S. and Goldstein, J.L. (1986) Science 232, 34-47] are known to play an important role in the homeostasis of lipoprotein metabolism *in vivo*, recognizing apo-B-100 and apo-E and taking primarily LDL as their ligands. Also, LRP, which is a macroprotein, has been found to primarily recognize apo-E and to take β-VLDL or chylomicron remnant as
40   a ligand. Moreover, it has been recently reported that LRP takes an α_2-macroglobulin/protease complex or a plasminogen activator/plasminogen activator inhibitor-1 complex as a ligand, and that LRP is a protein identical to the α_2-macroglobulin receptor. When these findings are taken together, LRP is considered to have a wide variety of functions in living bodies [Herz, J., Hamann, U., Rogne, S., Myklebost, O., Gausepohl, H. and Stanley, K.K. (1989) EMBO J. 7(13), 4119-4127; Brown, M.S., Herz, J., Kowal, R.C. and Goldstein, J.L. (1991) Current Opinion in Lipidology 2, 65-72; Herz,
45   J. (1993) Current Opinion in Lipidology 4, 107-113]. The gp 330, which was first identified as an antigen inducing rat Heymann nephritis, has been reported to have a ligand-binding capacity similar to that possessed by CRP α_2-macroglobulin receptor [Raychowdhury, R., Niles, J.L., McCluskey, R.T. and Smith, J.A. (1989) Science 244, 1163-1165; Pietromonaco, S., Kerjaschki, D., Binder, S., Ullrich, R. and Farquhar, G. (1990) Proc. Natl. Acad. Sci. U.S.A. 87, 1811-1815]. In addition, recently discovered VLDL receptors, which are found to take VLDL as a ligand, are considered to
50   have new functions including fatty acid metabolism, because they are predominantly found in tissues of the heart and muscles though they are rarely found in the liver [Takahashi, S., Kawarabayashi, Y., Nakai, T., Sakai, J. and Yamamoto, T. (1992) Proc. Natl. Acad. Sci. USA 89, 9252-9256].
Functions of these newly found receptors as lipoprotein receptors have been gradually elucidated through detailed *in vitro* analyses. However, significance of respective receptors in living bodies has mostly been left unknown. In addi-
55   tion, relations to remnant receptors, HDL receptors, etc., which have conventionally been identified or suggested by biochemical techniques, remain unknown. Presently, it is considered that these newly found receptors are products of genes different from those of the latter receptors. Thus, more lipoprotein receptors than originally guessed have become considered to participate in lipoprotein uptake into cells while interacting with each other to thereby function to maintain homeostasis of lipid metabolism in living bodies. However, from structural analyses of the genes of the afore-

mentioned newly-identified receptors, it is predicted that the genes of these receptors that take lipoproteins as ligands are developed from the same gene from which LDL receptors was developed, and thus they are within the same genetic family. This suggests that lipoprotein receptors that have conventionally been proposed may have structures similar to those of LDL receptors.

5 Accordingly, an object of the present invention is to provide the gene of a novel receptor in the LDL receptor family, as well as a protein coded by the gene.

The present inventors conducted careful studies so as to attain the above object, and found that by using part of rabbit LDL receptor cDNA as a probe there can be obtained a DNA fragment coding for a peptide having a structure similar to that of LDL receptors. Moreover, when using part of the obtained cDNA as a probe, a cDNA fragment having

10 a sequence similar to that of the cDNA can be obtained from the human tissue cDNA library. The present invention was accomplished based on these findings.

Summary of the Invention

15 The present invention provides DNA having a nucleotide sequence shown by Sequence ID No. 1 or No. 5; an LDL receptor analog protein having an amino acid sequence coded by the DNA; a recombinant vector comprising the DNA and a replicable vector; transformant cells which harbor the recombinant vector; and a method for the production of the LDL receptor analog protein.

20 Description of Preferred Embodiment

The cDNA of the present invention may be prepared, for example, by the following process.

Briefly, the process includes the following steps. (1) Through the use of rabbit LDL receptor cDNA as a probe, positive clones are screened out of a rabbit liver cDNA library. (2) Recombinant DNA is prepared using the separated pos-

25 itive clones, and a cDNA fragment is cut out of the resultant recombinant DNA through a treatment using a restriction enzyme The cDNA fragment is integrated into a plasmid vector. (3) Host cells are transformed using the obtained cDNA recombinant vector to thereby obtain transformant cells of the present invention. The obtained transformant cells are incubated so as to obtain a recombinant vector containing a DNA fragment of the present invention. The nucleotide sequence of the DNA fragment of the present invention contained in the resultant recombinant vector is determined. (4)

30 In tissue of a living body, there is detected expression of mRNA indicated by the nucleotide sequence of the cDNA of the present invention by using RNA blot hybridization method. (5) Through use of a rabbit cDNA fragment as a probe, positive clones are screened out of a human tissue cDNA library, and the nucleotide sequence of the clones is determined. (6) A recombinant vector for expression is prepared using the cDNA of the present invention. Through use of the thus-obtained vector, host cells are transformed to thereby obtain the transformants of the present invention. (7) Lig-

35 ands that are bound to protein expressed by the obtained transformants are detected by ligand blotting.

Each of the above-described steps will next be described.

(1) Screening for positive clones from a rabbit liver cDNA library:

40 A cDNA library may be prepared by the use of mRNA obtained from rabbit liver, reverse transcriptase, and a suitable vector, e.g., commercially available λgt10 vector.

A cDNA library thus prepared using λgt10 as a vector is subjected to a screening for positive clones by the application of a DNA hybridization method employing a cDNA probe, to thereby separate positive clones [Sambrook, J., Fritsch, E.F. and Maniatis, T. (1989) In: Molecular Cloning: A Laboratory Manual, pp 9.47-9.58, Cold Spring Harbor Lab-

45 oratory Press].

An exemplary cDNA which may be used as a probe is rabbit LDL receptor cDNA. Positive clones may be detected by autoradiography employing a DNA probe labelled with a radioisotope (^{32}P).

(2) Preparation of a cDNA recombinant vector:

50 Recombinant vector λgt10 phage DNA is extracted from the isolated positive clones and purified. The resultant purified recombinant vector λgt10 phage DNA is digested with a restriction enzyme EcoRI, to thereby separate a cDNA fragment from the vector DNA. The obtained cDNA fragment is integrated with a plasmid vector for cloning that has been similarly digested with EcoRI, thereby obtaining a recombinant plasmid vector. An exemplary plasmid vector

55 which may be used is pBluescript II.

(3) Recombinant vector, transformation of host cells using the recombinant vector, and preparation of DNA:

The obtained cDNA recombinant vector is introduced into a variety of host cells that are capable of utilizing the

genetic marker possessed by the recombinant vector, to thereby transform the host cells. Host cells are not particularly limited, with *E. coli* being preferred. For example, a variety of variants of the *E. coli* K12 strain, e.g., HB-101, may be used. In order to introduce the recombinant vector into host cells, a competent cell method may be used in combination with a treatment with calcium.

5   The thus-obtained transformant cells are cultured in a selective medium in accordance with the genetic marker of the vector. The recombinant vector of the present invention is collected from the cultured cells. The DNA nucleotide sequence of the cDNA contained in the obtained recombinant vector can be determined through use of a dideoxy sequence method [Sanger, F., Nicklen, S. and Coulson, A.R. (1977) Proc. Natl. Acd. Sci. USA 74, 5463-5467].

10   (4) RNA blot hybridization:

The expression in tissue of mRNA, indicated by the nucleotide sequence of the cDNA of the present invention, is detected using RNA blot hybridization.
First, mRNA is prepared using rabbit tissue. Commercially available oligo(dT)cellulose column may be used for the
15   preparation. In order to prepare mRNA from human tissue, there may be used a commercially available nylon membrane on which tissue poly(A)$^+$RNA from a variety of sources is present.
An exemplary probe is the rabbit cDNA obtained in the above-described step (3). mRNA may be detected by autoradiography employing a DNA probe labelled with a radioisotope (^{32}P).

20   (5) Screening of human tissue cDNA library for positive clones, and determination of nucleotide sequence:

An exemplary human tissue cDNA library which may be used is a commercially available human brain cDNA library.
Screening and nucleotide sequencing of the human brain cDNA library may be performed using a fragment of rab-
25   bit cDNA of the present invention as a probe in a manner similar to that used for the aforementioned rabbit liver cDNA library.

(6) Preparation of a recombinant vector for expression and transformation of host cells using the recombinant vector for expression:
30

In order to prepare an LDL receptor analog protein through use of cDNA of the present invention, the obtained cDNA and a vector for expression are first bonded to each other to thereby create a recombinant vector for expression. Vectors for expression which may be used for bonding are not particularly limited. For example, pBK-CMV may be used.
Host cells are transformed using the thus-obtained recombinant vector for expression, to thereby obtain a trans-
35   formant cell of the present invention. The obtained transformant cell is cultured so as to obtain cells that are capable of expressing the protein of the invention. Host cells are not particularly limited. For example, CHO cells may be used. In order to introduce the recombinant vector for expression into host cells, a calcium phosphate method may be used.
The thus-prepared transformant cells are incubated in a selective medium in accordance with the genetic marker of the vector, so as to express the LDL receptor analog protein of the present invention.
40

(7) Ligand analysis of the protein by ligand blotting:

After the resultant transformant cells are incubated, the expressed LDL receptor analog protein is solubilized using a solubilizer, e.g., Triton X-100, to thereby obtain a membrane protein fraction. The fraction is separated using SDS-
45   PAGE, and transferred onto, for example, a nitrocellulose membrane. Using a radio-labelled (^{125}I) lipoprotein as a probe, the analog protein can be detected by autoradiography. Exemplary lipoproteins which may be used include β-VLDL and LDL.

Examples:
50

The present invention will next be described in detail by way of example, which should not be construed as limiting the invention.

Example 1:
55

Preparation of a rabbit liver cDNA library:

From tissue of the liver of a male Japanese white rabbit, intact RNA was extracted through a guanidium thiocyanate/cesium chloride method. The obtained intact RNA was subjected to an oligo (dT) cellulose column method to

thereby obtain purified poly(A)$^+$RNA.

cDNA was synthesized in accordance with a method of Gubler and Hoffman [Gubler, U. and Hoffman, B.J. (1983) Gene 25, 263]. Briefly, cDNA was synthesized employing rabbit liver poly(A)$^+$RNA (as a template), a random primer, and moloney murine leukemia virus reverse transcriptase. The synthesized cDNA was transformed into double-stranded DNA using DNA polymerase I, and then subjected to an EcoRI methylase treatment. By the use of T4 DNA polymerase, the DNA was blunt-ended. The blunt-ended DNA was ligated to phosphorylated EcoRI linker pd (CCGAATTCGG) using a T4 DNA ligase, and the resultant ligated product was subjected to an additional digestion with EcoRI. cDNA fragments having a size not less than 1 kb were selected by agarose gel electrophoresis, and integrated into the EcoRI-digested site of λgt10 phage DNA using a T4 DNA ligase. The phage DNA was packaged in vitro, to thereby establish a rabbit liver cDNA library.

Example 2:

Cloning of cDNA of receptors in the rabbit LDL receptor family:

The cDNA library (1,000,000 plaques) prepared in Example 1 was subjected to screening using a plaque hybridization method and employing as a probe a segment of the cDNA obtained from a ligand binding region, the functional region, of the rabbit LDL receptor. Hybridization was performed at 42°C using 5 x SSC, 30% formamide, 1% SDS, 5 x Denhardt's, and 100 µg/ml salmon sperm DNA (ssDNA), followed by washing with 0.3 x SSC/0.1% SDS at 48°C. As a result, several positive clones were obtained. These cDNA clones were separated by performing this plaque hybridization method in a plurality of times. Subsequently, a cDNA fragment of each phage was subcloned into a plasmid vector pBluescript II, and the nucleotide sequence was analyzed using a dideoxy sequence method [Sanger, F., Nicklen, S. and Coulson, A.R. (1977) Proc. Natl. Acd. Sci. USA 74, 5463-5467]. Based on a putative amino acid sequence, LDL receptors themselves were excluded, and cDNA clones having a sequence very similar to that of LDL receptors were identified. Using these clones as cDNA probes, the cDNA library was screened to thereby obtain overlapping two clones. These were employed as new probes and similar procedure was performed, so as to obtain 5 cDNA clones. The DNA nucleotide sequence determined by these cDNA clones are shown as Sequence ID No. 3. The total length of the sequence was 6961 bp. In the open reading frame of 6639 bp (Sequence ID No. 1) which contained a sequence exhibiting high homology with LDL receptors, there existed on the 5' side an ATG codon which was presumably a translation initiating site and a successive highly hydrophobic sequence consisting of about 30 amino acids. Accordingly, the obtained cDNA was considered to contain the entirety of its length. A putative amino acid sequence is shown as Sequence ID No. 2. The protein consisted of 2213 amino acids. Comparison of the amino acid sequence of the protein with other amino acid sequence data registered at the Genebank, there was a very high similarity to LDL receptors. That is, amino acids 700 - 1,100 in the sequence were very similar to the EGF precursor homology region of LDL receptors, and amino acids 1,100 - 1,640 were also very similar to the ligand binding region of LDL receptors. When the amino acid sequence of the subject protein was compared with other lipoprotein receptor LRP, gp330, and VLDL receptors, similarity was not as high as that observed for LDL receptors. On the C-terminal side of the amino acid sequence of the protein, there was found a highly hydrophobic region which was very similar to the transmembrane region of LDL receptors.

Example 3:

From liver tissue and brain tissue of a male Japanese white rabbit, intact RNA was extracted through a guanidium thiocyanate/cesium chloride method. The obtained intact RNA was subjected to an oligo (dT) cellulose column method to thereby obtain purified poly(A)$^+$RNA. The poly(A)$^+$RNA specimens (10 µg each) was modified via a glyoxal method, electrophoresed on 1% agarose gel, and transferred onto a nylon membrane.

For human tissue mRNA, commercially available nylon membranes blotted with human tissue poly(A)$^+$RNA from various sources were used.

Using as a probe part of a ^{32}P-labelled rabbit cDNA of the present invention, hybridization was performed at 42°C using 50% (rabbit) or 40% (human) formamide, 0.1% SDS, 50 mM phosphate buffer, 5 x Denhardt's, 5 x SSC, and 200 µg/ml of ssDNA, followed by washing with 0.1 x SSC and 0.1% SDS at 50°C. Autoradiography was performed at -70°C for 2 days in the presence of intensifying screen. As a result, in both rabbit liver tissue and brain tissue, mRNA of about 7 kb was detected as well as mRNA of about 15 kb which was considered to result from alternative splicing or polyadenylation. The size of the mRNA of about 7 kb coincided with that of the rabbit cDNA of the present invention. Also, in human liver tissue and brain tissue, it was confirmed that mRNA having the same size was expressed.

Example 4:

Screening of human brain cDNA library for positive clones and determination of the nucleotide sequence of cDNA fragments

5

The human brain cDNA library used in this Example was a commercially obtained cDNA library which was constructed using λgt10 as a vector. Using partial cDNA of the present invention as a probe, screening of the cDNA library (300,000 plaques) was performed using a plaque hybridization method. Procedures of screening, cloning, and sequencing were as described in Example 2 of the present invention.

10
As a result of screening of the human brain cDNA library, positive clones containing a DNA fragment of about 3 kb were obtained. Analysis of the nucleotide sequence of part of the cDNA fragment revealed that the fragment was highly homologous to the cDNA of the present invention (Sequence ID No. 4).

Example 5:

15

Cloning of cDNA of receptors in the human LDL receptor family:

A human brain cDNA library was subjected to screening using fragments of the cDNA of the present invention and fragments of the cDNA obtained in Example 4 as probes. Procedures of screening, cloning, and sequencing were as
20 described in Example 2 of the present invention.

Through screening of the human brain cDNA library, two positive clones containing cDNA fragments of about 6 kb and about 3 kb were obtained. When their nucleotide sequence was analyzed, they were identified to be a cDNA clone containing the cDNA nucleotide sequence obtained in Example 4 and a cDNA clone that overlapped therewith. Using part of these cDNAs as probes, procedures similar to those as described above were performed, to thereby obtain
25 another cDNA clone The DNA nucleotide sequence indicated by these cDNA clones are shown as Sequence ID No. 7. The total length of the sequence was 6,843 bp. There was an open reading frame having a size of 6,642 bp (Sequence ID No. 5). A putative amino acid sequence is shown as Sequence ID No. 6. The protein consisted of 2,214 amino acids. Comparison of the amino acid sequence with that of rabbit protein shown by Sequence ID No. 2 revealed high homology of not less than 94%.

30

Example 6:

Creation of cells that express receptors in the rabbit LDL receptor family:

35
The cDNA as shown by Sequence ID No. 3 was ligated to phosphorylated EcoRI linker pd (CCGAATTCGG) by the use of a T4 DNA ligase, and the resultant ligated product was digested with EcoRI. Separately, a vector for expression, pBK-CMV was digested with EcoRI. The aforementioned DNA was ligated to the EcoRI-digested site of the vector using a T4 DNA ligase.

Using the resultant recombinant expression vector in a calcium phosphate method [Chen, C. and H. Okayama
40 (1987) Mol. Cell. Biol. 7, 2745-2752], host cells (CHO-1d1A7) were transformed. The resultant transformants were incubated in a Ham's F-12 selective medium supplemented with 500 μg/ml of G418, and viable cells were separated as LDL receptor analog protein-expressing cells. The cells were incubated further in the aforementioned medium.

Example 7:

45

Ligand analysis of the LDL receptor analog protein by ligand blotting:

The obtained LDL receptor analog protein-expressing cells and control cells were suspended in a buffer solution containing 200 mM Tris-maleic acid (pH 6.5), 2 mM calcium chloride, 0.5 mM PMSF, 2.5 μM leupeptin, and 1% Triton
50 X-100, to thereby solubilize the membrane protein. Solubilized membrane protein fractions were obtained through centrifugation, and electrophoresed by a 4.5-18% gradient SDS-PAGE. Thereafter, the protein was transferred onto a nitrocellulose membrane.

Incubation was performed in a buffer of 50 mM Tris-HCl (pH 8.0) containing ^{125}I-labelled β-VLDL (10 μg/ml), 2 mM calcium chloride, and 5% bovine serum albumin. Autoradiography was performed at room temperature.

55
A single band of about 250 kDa was detected in membrane protein fractions prepared using the present protein-expressing cells. This size coincided well with the molecular weight of 248 kDa calculated regarding the amino acid sequence (Sequence ID No. 2) deduced from the cDNA of the present invention. Although a similar band was detected for control cells, the expression level was much lower as compared with the case of the present protein-expressing cells.

Since the protein coded by the cDNA of the present invention is considered to be a novel LDL receptor family recep-

tor, it is expected that through analyses of this protein, details of lipoprotein metabolism mediated by the membrane receptor will be elucidated, and pathology of abnormal lipid metabolism which triggers onset and progress of arteriosclerosis will be clarified.

Sequence ID No. 1

Length of the Sequence: 6639

Type: nucleic acid

Strandedness: double

Topology: linear

Molecular type: cDNA to mRNA

Sequence:

```
ATGGCGACAC GGAGCAGCAG GAGGGAGTCG CGACTCCCCT TCCTATTCAC CCTGGTCGCG    60

CTGCTGCCGC CCGGGGCTCT CTGCGAGGTG TGGACGCGGA CACTGCACGG CGGCCGCGCG    120

CCCTTACCCC AGGAGCGGGG CTTCCGCGTG GTGCAGGGCG ACCCGCGCGA GCTGCGGCTG    180

TGGGAGCGCG GGGATGCCAG GGGGGCGAGC CGGGCGGACG AGAAGCCGCT CCGGAGGAGA    240

CGGAGCGCTG CCCTGCAGCC CGAGCCCATC AAGGTGTACG GACAGGTCAG CCTCAATGAT    300

TCCCACAATC AGATGGTGGT GCACTGGGCC GGAGAGAAAA GCAACGTGAT CGTGGCCTTG    360

GCCCGGGACA GCCTGGCGTT GGCCAGGCCC AGGAGCAGTG ATGTGTACGT GTCTTATGAC    420

TATGGAAAAT CATTCAATAA GATTTCAGAG AAATTGAACT TCGGCGCGGG AAATAACACA    480

GAGGCTGTGG TGGCCCAGTT CTACCACAGC CCTGCGGACA ACAAACGGTA CATCTTCGCA    540

GATGCCTACG CCCAGTATCT CTGGATCACG TTTGACTTCT GCAACACCAT CCATGGCTTT    600

TCCATCCCGT TCCGGGCAGC TGATCTCCTA CTCCACAGTA AGGCCTCCAA CCTTCTCCTG    660

CGCTTCGACA GGTCTCACCC CAACAAGCAG CTGTGGAAGT CGGATGATTT TGGCCAGACC    720

TGGATCATGA TTCAAGAACA CGTGAAGTCC TTTTCTTGGG GAATTGATCC CTATGACAAA    780

CCAAACACCA TCTACATCGA ACGGCACGAA CCTTCTGGCT ACTCCACGGT TTTCCGAAGT    840

ACAGACTTCT TCCAGTCCCG GGAAAACCAG GAAGTGATCT TGGAGGAAGT GAGAGACTTT    900

CAGCTTCGGG ACAAGTACAT GTTTGCTACA AAGGTGGTGC ATCTCTTGGG CAGTCCACTG    960

CAGTCTTCTG TCCAGCTCTG GGTCTCCTTT GGCCGGAAGC CCATGCGGGC CGCCCAGTTT    1020

GTTACAAGAC ATCCTATCAA CGAATATTAC ATCGCGGATG CCTCGGAGGA CCAGGTGTTT    1080

GTGTGTGTCA GTCACAGCAA CAACCGCACC AACCTCTACA TCTCGGAGGC AGAGGGCTTG    1140

AAGTTCTCTC TGTCCCTGGA GAACGTGCTC TACTACACCC CGGGAGGGGC CGGCAGTGAC    1200

ACCTTGGTGA GGTACTTTGC AAATGAACCG TTTGCTGACT TCCATCGTGT GGAAGGGTTG    1260
```

```
CAGGGAGTCT ACATTGCTAC TCTGATTAAT GGTTCTATGA ATGAGGAGAA CATGAGATCT   1320

GTCATCACCT TTGACAAAGG GGGCACCTGG GAATTTCTGC AGGCTCCAGC CTTCACGGGG   1380

TATGGAGAGA AAATCAACTG TGAGCTGTCC GAGGGCTGTT CCCTCCACCT GGCCCAGCGC   1440

CTCAGCCAGC TGCTCAACCT CCAGCTCCGG AGGATGCCCA TCCTGTCCAA GGAGTCGGCG   1500

CCTGGCCTCA TCATTGCCAC GGGCTCAGTG GGAAAGAACT TGGCTAGCAA GACAAACGTG   1560

TACATCTCTA GCAGTGCTGG AGCCAGGTGG CGAGAGGCAC TTCCTGGACC TCACTACTAT   1620

ACATGGGGAG ACCATGGCGG CATCATCATG GCCATTGCCC AAGGCATGGA AACCAACGAA   1680

CTGAAGTACA GTACCAACGA AGGGGAGACC TGGAAAGCCT TCACCTTCTC TGAGAAGCCC   1740

GTGTTTGTGT ATGGGCTCCT CACGGAACCC GGCGAGAAGA GCACGGTCTT CACCATCTTT   1800

GGCTCCAACA AGGAGAACGT GCACAGCTGG CTCATCCTCC AGGTCAATGC CACAGACGCC   1860

CTGGGGGTTC CTTGCACAGA GAACGACTAC AAGCTCTGGT CACCATCTGA TGAGCGGGGG   1920

AATGAGTGTT TGCTTGGACA CAAGACTGTT TTCAAACGGA GGACCCCGCA CGCCACATGC   1980

TTTAACGGAG AAGACTTTGA CAGGCCGGTG GTTGTGTCCA ACTGCTCCTG CACCCGGGAG   2040

GACTATGAGT GTGACTTTGG CTTCCGGATG AGTGAAGACT TGGCATTAGA GGTGTGTGTT   2100

CCAGATCCAG GATTTTCTGG AAAGTCCTCC CCTCCAGTGC CTTGTCCCGT GGGCTCTACG   2160

TACAGGCGAT CAAGAGGCTA CCGGAAGATT TCTGGGGACA CCTGTAGTGG AGGAGATGTT   2220

GAGGCACGGC TAGAAGGAGA GCTGGTCCCC TGTCCCCTGG CAGAAGAGAA CGAGTTCATC   2280

CTGTACGCCA CGCGCAAGTC CATCCACCGC TATGACCTGG CTTCCGGAAC CACGGAGCAG   2340

TTGCCCCTCA CTGGGTTGCG GGCAGCAGTG GCCCTGGACT TTGACTATGA GCACAACTGC   2400

CTGTATTGGT CTGACCTGGC CTTGGACGTC ATCCAGCGCC TCTGTTTGAA CGGGAGTACA   2460

GGACAAGAGG TGATCATCAA CTCTGACCTG GAGACGGTAG AAGCTTTGGC TTTTGAACCC   2520

CTCAGCCAAT TACTTTACTG GGTGGACGCA GGCTTTAAAA AGATCGAGGT AGCCAATCCA   2580

GATGGTGACT TCCGACTCAC CGTCGTCAAT TCCTCGGTGC TGGATCGGCC CCGGGCCCTG   2640

GTCCTTGTGC CCCAAGAAGG GATCATGTTC TGGACCGACT GGGGAGACCT GAAGCCTGGG   2700

ATTTATCGGA GCAACATGGA CGGATCTGCC GCCTATCGCC TCGTGTCGGA GGATGTGAAG   2760

TGGCCCAATG GCATTTCCGT GGACGATCAG TGGATCTACT GGACGGATGC CTACCTGGAC   2820

TGCATTGAGC GCATCACGTT CAGCGGCCAG CAGCGCTCCG TCATCCTGGA CAGACTCCCG   2880

CACCCCTATG CCATTGCTGT CTTTAAGAAT GAGATTTACT GGGATGACTG GTCACAGCTC   2940
```

```
AGCATATTCC GAGCTTCTAA GTACAGCGGG TCCCAGATGG AGATTCTGGC CAGCCAGCTC   3000
ACGGGGCTGA TGGACATGAA GATCTTCTAC AAGGGGAAGA ACACAGGAAG CAATGCGTGT   3060
GTACCCAGGC CGTGCAGCCT GCTGTGCCTG CCCAGAGCCA ACAACAGCAA AAGCTGCAGG   3120
TGTCCAGATG GCGTGGCCAG CAGTGTCCTC CCTTCCGGGG ACCTGATGTG TGACTGCCCT   3180
AAGGGCTACG AGCTGAAGAA CAACACGTGT GTCAAAGAAG AAGACACCTG TCTGCGCAAC   3240
CAGTACCGCT GCAGCAACGG GAACTGCATC AACAGCATCT GGTGGTGCGA TTTCGACAAC   3300
GACTGCGGAG ACATGAGCGA CGAGAAGAAC TGCCCTACCA CCATCTGCGA CCTGGACACC   3360
CAGTTCCGTT GCCAGGAGTC TGGGACGTGC ATCCCGCTCT CCTACAAATG TGACCTCGAG   3420
GATGACTGTG GGGACAACAG TGACGAAAGG CACTGTGAAA TGCACCAGTG CCGGAGCGAC   3480

GAATACAACT GCAGCTCGGG CATGTGCATC CGCTCCTCCT GGGTGTGCGA CGGGGACAAC   3540
GACTGCAGGG ACTGGTCCGA CGAGGCCAAC TGCACAGCCA TCTATCACAC CTGTGAGGCC   3600
TCCAACTTCC AGTGCCGCAA CGGGCACTGC ATCCCCCAGC GGTGGGCGTG TGACGGCGAC   3660
GCCGACTGCC AGGATGGCTC TGATGAGGAT CCAGCCAACT GTGAGAAGAA GTGCAACGGC   3720
TTCCGCTGCC CGAACGGCAC CTGCATTCCC TCCACCAAGC ACTGTGACGG CCTGCACGAT   3780
TGCTCGGACG GCTCCGACGA GCAGCACTGC GAGCCCCTGT GTACACGGTT CATGGACTTC   3840
GTGTGTAAGA ACCGCCAGCA GTGCCTCTTC CACTCCATGG TGTGCGATGG GATCATCCAG   3900
TGCCGTGACG GCTCCGACGA GGACCCAGCC TTTGCAGGAT GCTCCCGAGA CCCCGAGTTC   3960
CACAAGGTGT GCGATGAGTT CGGCTTCCAG TGTCAGAACG GCGTGTGCAT CAGCTTGATC   4020
TGGAAGTGCG ACGGGATGGA TGACTGCGGG GACTACTCCG ACGAGGCCAA CTGTGAAAAC   4080
CCCACAGAAG CCCCCAACTG CTCCCGCTAC TTCCAGTTCC GGTGTGACAA TGGCCACTGC   4140
ATCCCCAACA GGTGGAAGTG TGACAGGGAG AATGACTGTG GGGACTGGTC CGACGAGAAG   4200
GACTGTGGAG ATTCACATGT ACTTCCGTCT ACGACTCCTG CACCCTCCAC GTGTCTGCCC   4260
AATTACTACC GCTGCGGCGG GGGGGCCTGC GTGATAGACA CGTGGGTTTG TGACGGGTAC   4320
CGAGATTGCG CAGATGGATC CGACGAGGAA GCCTGCCCCT CGCTCCCCAA TGTCACTGCC   4380
ACCTCCTCCC CCTCCCAGCC TGGACGATGC GACCGATTTG AGTTTGAGTG CCACCAGCCA   4440
AAGAAGTGCA TCCCTAACTG GAGACGCTGT GACGGCCATC AGGATTGCCA GGATGGCCAG   4500
GACGAGGCCA ACTGCCCCAC TCACAGCACC TTGACCTGCA TGAGCTGGGA GTTCAAGTGT   4560
GAGGATGGCG AGGCCTGCAT CGTGCTGTCA GAACGCTGCG ACGGCTTCCT GGACTGCTCA   4620
```

```
GATGAGAGCG ACGAGAAGGC CTGCAGTGAT GAGTTAACTG TATACAAAGT ACAGAATCTT   4680
CAGTGGACAG CTGACTTCTC TGGGAATGTC ACTTTGACCT GGATGCGGCC CAAAAAAATG   4740
CCCTCTGCTG CTTGTGTATA CAACGTGTAC TATAGAGTTG TTGGAGAGAG CATATGGAAG   4800
ACTCTGGAGA CTCACAGCAA TAAGACAAAC ACTGTATTAA AAGTGTTGAA ACCAGATACC   4860
ACCTACCAGG TTAAAGTGCA GGTTCAGTGC CTGAGCAAGG TGCACAACAC CAATGACTTT   4920
GTGACCTTGA GAACTCCAGA GGGATTGCCA GACGCCCCTC AGAACCTCCA GCTGTCGCTC   4980
CACGGGGAAG AGGAAGGTGT GATTGTGGGC CACTGGAGCC CTCCCACCCA CACCCACGGC   5040
CTCATTCGCG AATACATTGT AGAGTATAGC AGGAGTGGTT CCAAGGTGTG GACTTCAGAA   5100
AGGGCTGCTA GTAACTTTAC AGAAATAAAG AACTTGTTGG TCAACACCCT GTACACCGTC   5160
AGAGTGGCTG CGGTGACGAG TCGTGGGATA GGAAACTGGA GCGATTCCAA ATCCATTACC   5220
ACCGTGAAAG GAAAAGCGAT CCCGCCACCA AATATCCACA TTGACAACTA CGATGAAAAT   5280
TCCCTGAGTT TTACCCTGAC CGTGGATGGG AACATCAAGG TGAATGGCTA TGTGGTGAAC   5340
CTTTTCTGGG CATTTGACAC CCACAAACAA GAGAAGAAAA CCATGAACTT CCAAGGGAGC   5400
TCAGTGTCCC ACAAAGTTGG CAATCTGACA GCACAGACGG CCTATGAGAT TTCCGCCTGG   5460
GCCAAGACTG ACTTGGGCGA TAGTCCTCTG TCATTTGAGC ATGTCACGAC CAGAGGGGTT   5520
CGCCCACCTG CTCCTAGCCT CAAGGCCAGG GCTATCAATC AGACTGCAGT GGAATGCACC   5580
TGGACAGGCC CCAGGAATGT GGTGTATGGC ATTTTCTATG CCACATCCTT CCTGGACCTC   5640
TACCGCAACC CAAGCAGCCT GACCACGCCG CTGCACAACG CAACCGTGCT CGTCGGTAAG   5700
GATGAGCAGT ATCTGTTTCT GGTCCGGGTG GTGATGCCCT ACCAAGGGCC GTCCTCGGAC   5760
TACGTGGTCG TGAAGATGAT CCCGGACAGC AGGCTTCCTC CCCGGCACCT GCATGCCGTT   5820
CACACCGGCA AGACCTCGGC CGTCATCAAG TGGGAGTCGC CCTACGACTC TCCTGACCAG   5880
GACCTGTTCT ATGCGATCGC AGTTAAAGAT CTGATACGAA AGACGGACCG GAGCTACAAA   5940
GTCAAGTCCC GCAACAGCAC CGTGGAGTAC ACCCTGAGCA AGCTGGAGCC CGGAGGGAAA   6000
TACCACGTCA TTGTGCAGCT GGGGAACATG AGCAAAGATG CCAGTGTGAA GATCACCACC   6060
GTTTCGTTAT CGGCACCCGA TGCCTTAAAA ATCATAACAG AAAATGACCA CGTCCTTCTC   6120
TTCTGGAAAA GTCTAGCTCT AAAGGAAAAG TATTTTAACG AAAGCAGGGG CTACGAGATA   6180
CACATGTTTG ATAGCGCCAT GAATATCACC GCATACCTTG GGAATACTAC TGACAATTTC   6240
TTTAAAATTT CCAACCTGAA GATGGGTCAC AATTACACAT TCACGGTCCA GGCACGATGC   6300
```

```
CTTTTGGGCA GCCAGATCTG CGGGGAGCCT GCCGTGCTAC TGTATGATGA GCTGGGGTCT    6360
GGTGGCGATG CGTCGGCGAT GCAGGCTGCC AGGTCTACTG ATGTCGCCGC CGTGGTGGTG    6420
CCCATCCTGT TTCTGATACT GCTGAGCCTG GGGGTCGGGT TTGCCATCCT GTACACGAAG    6480
CATCGGAGGC TGCAGAGCAG CTTCACCGCC TTCGCCAACA GCCACTACAG CTCCAGACTC    6540
GGCTCCGCCA TCTTCTCCTC TGGGGATGAC TTGGGGGAGG ATGATGAACA TGCTCCTATG    6600
ATCACTGGAT TTTCGGACGA CGTCCCCATG GTGATAGCC                           6639
```

Sequence ID No. 2

Length of the Sequence: 2213

Type: amino acid

Topology: linear

Molecular type: Protein

Sequence:

```
Met Ala Thr Arg Ser Ser Arg Arg Glu Ser Arg Leu Pro Phe Leu Phe
                    5                 10                  15
Thr Leu Val Ala Leu Leu Pro Pro Gly Ala Leu Cys Glu Val Trp Thr
                20                  25                  30
Arg Thr Leu His Gly Gly Arg Ala Pro Leu Pro Gln Glu Arg Gly Phe
            35                  40                  45
Arg Val Val Gln Gly Asp Pro Arg Glu Leu Arg Leu Trp Glu Arg Gly
        50                  55                  60
Asp Ala Arg Gly Ala Ser Arg Ala Asp Glu Lys Pro Leu Arg Arg Arg
    65                  70                  75                  80
Arg Ser Ala Ala Leu Gln Pro Glu Pro Ile Lys Val Tyr Gly Gln Val
                85                  90                  95
Ser Leu Asn Asp Ser His Asn Gln Met Val Val His Trp Ala Gly Glu
                100                 105                 110
Lys Ser Asn Val Ile Val Ala Leu Ala Arg Asp Ser Leu Ala Leu Ala
            115                 120                 125
```

Arg Pro Arg Ser Ser Asp Val Tyr Val Ser Tyr Asp Tyr Gly Lys Ser
     130   135   140

Phe Asn Lys Ile Ser Glu Lys Leu Asn Phe Gly Ala Gly Asn Asn Thr
145   150   155   160

Glu Ala Val Val Ala Gln Phe Tyr His Ser Pro Ala Asp Asn Lys Arg
          165   170   175

Tyr Ile Phe Ala Asp Ala Tyr Ala Gln Tyr Leu Trp Ile Thr Phe Asp
          180   185   190

Phe Cys Asn Thr Ile His Gly Phe Ser Ile Pro Phe Arg Ala Ala Asp
          195   200   205

Leu Leu Leu His Ser Lys Ala Ser Asn Leu Leu Leu Gly Phe Asp Arg
      210   215   220

Ser His Pro Asn Lys Gln Leu Trp Lys Ser Asp Asp Phe Gly Gln Thr
225   230   235   240

Trp Ile Met Ile Gln Glu His Val Lys Ser Phe Ser Trp Gly Ile Asp
          245   250   255

Pro Tyr Asp Lys Pro Asn Thr Ile Tyr Ile Glu Arg His Glu Pro Ser
          260   265   270

Gly Tyr Ser Thr Val Phe Arg Ser Thr Asp Phe Phe Gln Ser Arg Glu
          275   280   285

Asn Gln Glu Val Ile Leu Glu Glu Val Arg Asp Phe Gln Leu Arg Asp
      290   295   300

Lys Tyr Met Phe Ala Thr Lys Val Val His Leu Leu Gly Ser Pro Leu
305   310   315   320

Gln Ser Ser Val Gln Leu Trp Val Ser Phe Gly Arg Lys Pro Met Arg
          325   330   335

Ala Ala Gln Phe Val Thr Arg His Pro Ile Asn Glu Tyr Tyr Ile Ala
          340   345   350

```
Asp Ala Ser Glu Asp Gln Val Phe Val Cys Val Ser His Ser Asn Asn
        355             360             365
Arg Thr Asn Leu Tyr Ile Ser Glu Ala Glu Gly Leu Lys Phe Ser Leu
        370             375             380
Ser Leu Glu Asn Val Leu Tyr Tyr Thr Pro Gly Gly Ala Gly Ser Asp
385             390             395             400
Thr Leu Val Arg Tyr Phe Ala Asn Glu Pro Phe Ala Asp Phe His Arg
                405             410             415
Val Glu Gly Leu Gln Gly Val Tyr Ile Ala Thr Leu Ile Asn Gly Ser
                420             425             430
Met Asn Glu Glu Asn Met Arg Ser Val Ile Thr Phe Asp Lys Gly Gly
        435             440             445
Thr Trp Glu Phe Leu Gln Ala Pro Ala Phe Thr Gly Tyr Gly Glu Lys
        450             455             460
Ile Asn Cys Glu Leu Ser Glu Gly Cys Ser Leu His Leu Ala Gln Arg
465             470             475             480
Leu Ser Gln Leu Leu Asn Leu Gln Leu Arg Arg Met Pro Ile Leu Ser
                485             490             495
Lys Glu Ser Ala Pro Gly Leu Ile Ile Ala Thr Gly Ser Val Gly Lys
                500             505             510
Asn Leu Ala Ser Lys Thr Asn Val Tyr Ile Ser Ser Ser Ala Gly Ala
        515             520             525
Arg Trp Arg Glu Ala Leu Pro Gly Pro His Tyr Tyr Thr Trp Gly Asp
        530             535             540
His Gly Gly Ile Ile Met Ala Ile Ala Gln Gly Met Glu Thr Asn Glu
545             550             555             560
Leu Lys Tyr Ser Thr Asn Glu Gly Glu Thr Trp Lys Ala Phe Thr Phe
                565             570             575
```

Ser Glu Lys Pro Val Phe Val Tyr Gly Leu Leu Thr Glu Pro Gly Glu
           580   585   590

Lys Ser Thr Val Phe Thr Ile Phe Gly Ser Asn Lys Glu Asn Val His
       595   600   605

Ser Trp Leu Ile Leu Gln Val Asn Ala Thr Asp Ala Leu Gly Val Pro
   610   615   620

Cys Thr Glu Asn Asp Tyr Lys Leu Trp Ser Pro Ser Asp Glu Arg Gly
625   630   635   640

Asn Glu Cys Leu Leu Gly His Lys Thr Val Phe Lys Arg Arg Thr Pro
         645   650   655

His Ala Thr Cys Phe Asn Gly Glu Asp Phe Asp Arg Pro Val Val Val
         660   665   670

Ser Asn Cys Ser Cys Thr Arg Glu Asp Tyr Glu Cys Asp Phe Gly Phe
         675   680   685

Arg Met Ser Glu Asp Leu Ala Leu Glu Val Cys Val Pro Asp Pro Gly
      690   695   700

Phe Ser Gly Lys Ser Ser Pro Pro Val Pro Cys Pro Val Gly Ser Thr
705   710   715   720

Tyr Arg Arg Ser Arg Gly Tyr Arg Lys Ile Ser Gly Asp Thr Cys Ser
         725   730   735

Gly Gly Asp Val Glu Ala Arg Leu Glu Gly Glu Leu Val Pro Cys Pro
         740   745   750

Leu Ala Glu Glu Asn Glu Phe Ile Leu Tyr Ala Thr Arg Lys Ser Ile
         755   760   765

His Arg Tyr Asp Leu Ala Ser Gly Thr Thr Glu Gln Leu Pro Leu Thr
         770   775   780

Gly Leu Arg Ala Ala Val Ala Leu Asp Phe Asp Tyr Glu His Asn Cys
785   790   795   800

Leu Tyr Trp Ser Asp Leu Ala Leu Asp Val Ile Gln Arg Leu Cys Leu
                    805   810   815

Asn Gly Ser Thr Gly Gln Glu Val Ile Ile Asn Ser Asp Leu Glu Thr
                820   825   830

Val Glu Ala Leu Ala Phe Glu Pro Leu Ser Gln Leu Leu Tyr Trp Val
               835   840   845

Asp Ala Gly Phe Lys Lys Ile Glu Val Ala Asn Pro Asp Gly Asp Phe
             850   855   860

Arg Leu Thr Val Val Asn Ser Ser Val Leu Asp Arg Pro Arg Ala Leu
865   870   875   880

Val Leu Val Pro Gln Glu Gly Ile Met Phe Trp Thr Asp Trp Gly Asp
                885   890   895

Leu Lys Pro Gly Ile Tyr Arg Ser Asn Met Asp Gly Ser Ala Ala Tyr
               900   905   910

Arg Leu Val Ser Glu Asp Val Lys Trp Pro Asn Gly Ile Ser Val Asp
               915   920   925

Asp Gln Trp Ile Tyr Trp Thr Asp Ala Tyr Leu Asp Cys Ile Glu Arg
         930   935   940

Ile Thr Phe Ser Gly Gln Gln Arg Ser Val Ile Leu Asp Arg Leu Pro
945   950   955   960

His Pro Tyr Ala Ile Ala Val Phe Lys Asn Glu Ile Tyr Trp Asp Asp
               965   970   975

Trp Ser Gln Leu Ser Ile Phe Arg Ala Ser Lys Tyr Ser Gly Ser Gln
               980   985   990

Met Glu Ile Leu Ala Ser Gln Leu Thr Gly Leu Met Asp Met Lys Ile
           995   1000   1005

Phe Tyr Lys Gly Lys Asn Thr Gly Ser Asn Ala Cys Val Pro Arg Pro
     1010   1015   1020

Cys Ser Leu Leu Cys Leu Pro Arg Ala Asn Asn Ser Lys Ser Cys Arg
1025   1030   1035   1040

Cys Pro Asp Gly Val Ala Ser Ser Val Leu Pro Ser Gly Asp Leu Met
    1045   1050   1055

Cys Asp Cys Pro Lys Gly Tyr Glu Leu Lys Asn Asn Thr Cys Val Lys
    1060   1065   1070

Glu Glu Asp Thr Cys Leu Arg Asn Gln Tyr Arg Cys Ser Asn Gly Asn
    1075   1080   1085

Cys Ile Asn Ser Ile Trp Trp Cys Asp Phe Asp Asn Asp Cys Gly Asp
    1090   1095   1100

Met Ser Asp Glu Lys Asn Cys Pro Thr Thr Ile Cys Asp Leu Asp Thr
1105   1110   1115   1120

Gln Phe Arg Cys Gln Glu Ser Gly Thr Cys Ile Pro Leu Ser Tyr Lys
    1125   1130   1135

Cys Asp Leu Glu Asp Asp Cys Gly Asp Asn Ser Asp Glu Arg His Cys
    1140   1145   1150

Glu Met His Gln Cys Arg Ser Asp Glu Tyr Asn Cys Ser Ser Gly Met
    1155   1160   1165

Cys Ile Arg Ser Ser Trp Val Cys Asp Gly Asp Asn Asp Cys Arg Asp
    1170   1175   1180

Trp Ser Asp Glu Ala Asn Cys Thr Ala Ile Tyr His Thr Cys Glu Ala
1185   1190   1195   1200

Ser Asn Phe Gln Cys Arg Asn Gly His Cys Ile Pro Gln Arg Trp Ala
    1205   1210   1215

Cys Asp Gly Asp Ala Asp Cys Gln Asp Gly Ser Asp Glu Asp Pro Ala
    1220   1225   1230

Asn Cys Glu Lys Lys Cys Asn Gly Phe Arg Cys Pro Asn Gly Thr Cys
    1235   1240   1245

Ile Pro Ser Thr Lys His Cys Asp Gly Leu His Asp Cys Ser Asp Gly
     1250   1255   1260

Ser Asp Glu Gln His Cys Glu Pro Leu Cys Thr Arg Phe Met Asp Phe
1265   1270   1275   1280

Val Cys Lys Asn Arg Gln Gln Cys Leu Phe His Ser Met Val Cys Asp
          1285   1290   1295

Gly Ile Ile Gln Cys Arg Asp Gly Ser Asp Glu Asp Pro Ala Phe Ala
        1300   1305   1310

Gly Cys Ser Arg Asp Pro Glu Phe His Lys Val Cys Asp Glu Phe Gly
     1315   1320   1325

Phe Gln Cys Gln Asn Gly Val Cys Ile Ser Leu Ile Trp Lys Cys Asp
     1330   1335   1340

Gly Met Asp Asp Cys Gly Asp Tyr Ser Asp Glu Ala Asn Cys Glu Asn
1345   1350   1355   1360

Pro Thr Glu Ala Pro Asn Cys Ser Arg Tyr Phe Gln Phe Arg Cys Asp
          1365   1370   1375

Asn Gly His Cys Ile Pro Asn Arg Trp Lys Cys Asp Arg Glu Asn Asp
        1380   1385   1390

Cys Gly Asp Trp Ser Asp Glu Lys Asp Cys Gly Asp Ser His Val Leu
     1395   1400   1405

Pro Ser Thr Thr Pro Ala Pro Ser Thr Cys Leu Pro Asn Tyr Tyr Arg
     1410   1415   1420

Cys Gly Gly Gly Ala Cys Val Ile Asp Thr Trp Val Cys Asp Gly Tyr
1425   1430   1435   1440

Arg Asp Cys Ala Asp Gly Ser Asp Glu Glu Ala Cys Pro Ser Leu Pro
        1445   1450   1455

Asn Val Thr Ala Thr Ser Ser Pro Ser Gln Pro Gly Arg Cys Asp Arg
        1460   1465   1470

Phe Glu Phe Glu Cys His Gln Pro Lys Lys Cys Ile Pro Asn Trp Arg
        1475   1480   1485

Arg Cys Asp Gly His Gln Asp Cys Gln Asp Gly Gln Asp Glu Ala Asn
        1490   1495   1500

Cys Pro Thr His Ser Thr Leu Thr Cys Met Ser Trp Glu Phe Lys Cys
1505   1510   1515   1520

Glu Asp Gly Glu Ala Cys Ile Val Leu Ser Glu Arg Cys Asp Gly Phe
                 1525   1530   1535

Leu Asp Cys Ser Asp Glu Ser Asp Glu Lys Ala Cys Ser Asp Glu Leu
        1540   1545   1550

Thr Val Tyr Lys Val Gln Asn Leu Gln Trp Thr Ala Asp Phe Ser Gly
        1555   1560   1565

Asn Val Thr Leu Thr Trp Met Arg Pro Lys Lys Met Pro Ser Ala Ala
        1570   1575   1580

Cys Val Tyr Asn Val Tyr Tyr Arg Val Val Gly Glu Ser Ile Trp Lys
1585   1590   1595   1600

Thr Leu Glu Thr His Ser Asn Lys Thr Asn Thr Val Leu Lys Val Leu
                 1605   1610   1615

Lys Pro Asp Thr Thr Tyr Gln Val Lys Val Gln Val Gln Cys Leu Ser
        1620   1625   1630

Lys Val His Asn Thr Asn Asp Phe Val Thr Leu Arg Thr Pro Glu Gly
        1635   1640   1645

Leu Pro Asp Ala Pro Gln Asn Leu Gln Leu Ser Leu His Gly Glu Glu
        1650   1655   1660

Glu Gly Val Ile Val Gly His Trp Ser Pro Pro Thr His Thr His Gly
1665   1670   1675   1680

Leu Ile Arg Glu Tyr Ile Val Glu Tyr Ser Arg Ser Gly Ser Lys Val
        1685   1690   1695

Trp Thr Ser Glu Arg Ala Ala Ser Asn Phe Thr Glu Ile Lys Asn Leu
.1700   1705   1710

Leu Val Asn Thr Leu Tyr Thr Val Arg Val Ala Ala Val Thr Ser Arg
   1715   1720   1725

Gly Ile Gly Asn Trp Ser Asp Ser Lys Ser Ile Thr Thr Val Lys Gly
   1730   1735   1740

Lys Ala Ile Pro Pro Pro Asn Ile His Ile Asp Asn Tyr Asp Glu Asn
1745   1750   1755   1760

Ser Leu Ser Phe Thr Leu Thr Val Asp Gly Asn Ile Lys Val Asn Gly
     1765   1770   1775

Tyr Val Val Asn Leu Phe Trp Ala Phe Asp Thr His Lys Gln Glu Lys
    1780   1785   1790

Lys Thr Met Asn Phe Gln Gly Ser Ser Val Ser His Lys Val Gly Asn
   1795   1800   1805

Leu Thr Ala Gln Thr Ala Tyr Glu Ile Ser Ala Trp Ala Lys Thr Asp
   1810   1815   1820

Leu Gly Asp Ser Pro Leu Ser Phe Glu His Val Thr Thr Arg Gly Val
1825   1830   1835   1840

Arg Pro Pro Ala Pro Ser Leu Lys Ala Arg Ala Ile Asn Gln Thr Ala
    1845   1850   1855

Val Glu Cys Thr Trp Thr Gly Pro Arg Asn Val Val Tyr Gly Ile Phe
    1860   1865   1870

Tyr Ala Thr Ser Phe Leu Asp Leu Tyr Arg Asn Pro Ser Ser Leu Thr
    1875   1880   1885

Thr Pro Leu His Asn Ala Thr Val Leu Val Gly Lys Asp Glu Gln Tyr
   1890   1895   1900

Leu Phe Leu Val Arg Val Val Met Pro Tyr Gln Gly Pro Ser Ser Asp
1905   1910   1915   1920

Tyr Val Val Val Lys Met Ile Pro Asp Ser Arg Leu Pro Pro Arg His
            1925   1930   1935

Leu His Ala Val His Thr Gly Lys Thr Ser Ala Val Ile Lys Trp Glu
        1940   1945   1950

Ser Pro Tyr Asp Ser Pro Asp Gln Asp Leu Phe Tyr Ala Ile Ala Val
        1955   1960   1965

Lys Asp Leu Ile Arg Lys Thr Asp Arg Ser Tyr Lys Val Lys Ser Arg
        1970   1975   1980

Asn Ser Thr Val Glu Tyr Thr Leu Ser Lys Leu Glu Pro Gly Gly Lys
1985   1990   1995   2000

Tyr His Val Ile Val Gln Leu Gly Asn Met Ser Lys Asp Ala Ser Val
        2005   2010   2015

Lys Ile Thr Thr Val Ser Leu Ser Ala Pro Asp Ala Leu Lys Ile Ile
        2020   2025   2030

Thr Glu Asn Asp His Val Leu Leu Phe Trp Lys Ser Leu Ala Leu Lys
        2035   2040   2045

Glu Lys Tyr Phe Asn Glu Ser Arg Gly Tyr Glu Ile His Met Phe Asp
        2050   2055   2060

Ser Ala Met Asn Ile Thr Ala Tyr Leu Gly Asn Thr Thr Asp Asn Phe
2065   2070   2075   2080

Phe Lys Ile Ser Asn Leu Lys Met Gly His Asn Tyr Thr Phe Thr Val
        2085   2090   2095

Gln Ala Arg Cys Leu Leu Gly Ser Gln Ile Cys Gly Glu Pro Ala Val
        2100   2105   2110

Leu Leu Tyr Asp Glu Leu Gly Ser Gly Gly Asp Ala Ser Ala Met Gln
        2115   2120   2125

Ala Ala Arg Ser Thr Asp Val Ala Ala Val Val Val Pro Ile Leu Phe
        2130   2135   2140

Leu Ile Leu Leu Ser Leu Gly Val Gly Phe Ala Ile Leu Tyr Thr Lys

2145   2150   2155   2160

His Arg Arg Leu Gln Ser Ser Phe Thr Ala Phe Ala Asn Ser His Tyr

          2165   2170   2175

Ser Ser Arg Leu Gly Ser Ala Ile Phe Ser Ser Gly Asp Asp Leu Gly

          2180   2185   2190

Glu Asp Asp Glu Asp Ala Pro Met Ile Thr Gly Phe Ser Asp Asp Val

          2195   2200   2205

Pro Met Val Ile Ala

          2210

Sequence ID No. 3

Length of the Sequence: 6961

Type: nucleic acid

Strandedness: double

Topology: linear

Molecular type: cDNA to mRNA

Feature:

 Name/Key: sig peptide

 Location: 178..261

 Identification method: S

 Name/Key: mat peptide

 Location: 262..6816

 Identification method: S

Sequence:

CCGCGAGCCG CACACGTGAC GGCGCCGCGC CGCGCCGCGC CGCGCCGAGC GGGACCCAGC 60

GGCTGCCCGG AGCCCCGGGA GCGGCGCGCG CGCGGCCCCG GCCCCGCCGC TCGGCCGGCG 120

GCGCGCTGCA CATTCTCTCC TGGCGGCGGC GCCACCTGCA GCCGCGTTCG CCCGAACATG 180

                                                              Met

21

1

```
GCG ACA CGG AGC AGC AGG AGG GAG TCG CGA CTC CCC TTC CTA TTC ACC     228
Ala Thr Arg Ser Ser Arg Arg Glu Ser Arg Leu Pro Phe Leu Phe Thr
              5                 10                15
CTG GTC GCG CTG CTG CCG CCC GGG GCT CTC TGC GAG GTG TGG ACG CGG     276
Leu Val Ala Leu Leu Pro Pro Gly Ala Leu Cys Glu Val Trp Thr Arg
              20                25                30
ACA CTG CAC GGC GGC CGC GCG CCC TTA CCC CAG GAG CGG GGC TTC CGC     324
Thr Leu His Gly Gly Arg Ala Pro Leu Pro Gln Glu Arg Gly Phe Arg
          35                40                45
GTG GTG CAG GGC GAC CCG CGC GAG CTG CGG CTG TGG GAG CGC GGG GAT     372
Val Val Gln Gly Asp Pro Arg Glu Leu Arg Leu Trp Glu Arg Gly Asp
50                55                60                65
GCC AGG GGG GCG AGC CGG GCG GAC GAG AAG CCG CTC CGG AGG AGA CGG     420
Ala Arg Gly Ala Ser Arg Ala Asp Glu Lys Pro Leu Arg Arg Arg Arg
                70                75                80
AGC GCT GCC CTG CAG CCC GAG CCC ATC AAG GTG TAC GGA CAG GTC AGC     468
Ser Ala Ala Leu Gln Pro Glu Pro Ile Lys Val Tyr Gly Gln Val Ser
                85                90                95
CTC AAT GAT TCC CAC AAT CAG ATG GTG GTG CAC TGG GCC GGA GAG AAA     516
Leu Asn Asp Ser His Asn Gln Met Val Val His Trp Ala Gly Glu Lys
              100               105               110
AGC AAC GTG ATC GTG GCC TTG GCC CGG GAC AGC CTG GCG TTG GCC AGG     564
Ser Asn Val Ile Val Ala Leu Ala Arg Asp Ser Leu Ala Leu Ala Arg
              115               120               125
CCC AGG AGC AGT GAT GTG TAC GTG TCT TAT GAC TAT GGA AAA TCA TTC     612
Pro Arg Ser Ser Asp Val Tyr Val Ser Tyr Asp Tyr Gly Lys Ser Phe
          130               135               140               145
```

AAT AAG ATT TCA GAG AAA TTG AAC TTC GGC GCG GGA AAT AAC ACA GAG   660
Asn Lys Ile Ser Glu Lys Leu Asn Phe Gly Ala Gly Asn Asn Thr Glu
                150   155   160

GCT GTG GTG GCC CAG TTC TAC CAC AGC CCT GCG GAC AAC AAA CGG TAC   708
Ala Val Val Ala Gln Phe Tyr His Ser Pro Ala Asp Asn Lys Arg Tyr
                165   170   175

ATC TTC GCA GAT GCC TAC GCC CAG TAT CTC TGG ATC ACG TTT GAC TTC   756
Ile Phe Ala Asp Ala Tyr Ala Gln Tyr Leu Trp Ile Thr Phe Asp Phe
                180   185   190

TGC AAC ACC ATC CAT GGC TTT TCC ATC CCG TTC CGG GCA GCT GAT CTC   804
Cys Asn Thr Ile His Gly Phe Ser Ile Pro Phe Arg Ala Ala Asp Leu
             195   200   205

CTA CTC CAC AGT AAG GCC TCC AAC CTT CTC CTG GGC TTC GAC AGG TCT   852
Leu Leu His Ser Lys Ala Ser Asn Leu Leu Leu Gly Phe Asp Arg Ser
210   215   220   225

CAC CCC AAC AAG CAG CTG TGG AAG TCG GAT GAT TTT GGC CAG ACC TGG   900
His Pro Asn Lys Gln Leu Trp Lys Ser Asp Asp Phe Gly Gln Thr Trp
                230   235   240

ATC ATG ATT CAA GAA CAC GTG AAG TCC TTT TCT TGG GGA ATT GAT CCC   948
Ile Met Ile Gln Glu His Val Lys Ser Phe Ser Trp Gly Ile Asp Pro
                245   250   255

TAT GAC AAA CCA AAC ACC ATC TAC ATC GAA CGG CAC GAA CCT TCT GGC   996
Tyr Asp Lys Pro Asn Thr Ile Tyr Ile Glu Arg His Glu Pro Ser Gly
               260   265   270

TAC TCC ACG GTT TTC CGA AGT ACA GAC TTC TTC CAG TCC CGG GAA AAC   1044
Tyr Ser Thr Val Phe Arg Ser Thr Asp Phe Phe Gln Ser Arg Glu Asn
             275   280   285

CAG GAA GTG ATC TTG GAG GAA GTG AGA GAC TTT CAG CTT CGG GAC AAG   1092

```
     Gln Glu Val Ile Leu Glu Glu Val Arg Asp Phe Gln Leu Arg Asp Lys
     290             295             300             305

     TAC ATG TTT GCT ACA AAG GTG GTG CAT CTC TTG GGC AGT CCA CTG CAG    1140
     Tyr Met Phe Ala Thr Lys Val Val His Leu Leu Gly Ser Pro Leu Gln
                 310             315             320

     TCT TCT GTC CAG CTC TGG GTC TCC TTT GGC CGG AAG CCC ATG CGG GCC    1188
     Ser Ser Val Gln Leu Trp Val Ser Phe Gly Arg Lys Pro Met Arg Ala
                 325             330             335

     GCC CAG TTT GTT ACA AGA CAT CCT ATC AAC GAA TAT TAC ATC GCG GAT    1236
     Ala Gln Phe Val Thr Arg His Pro Ile Asn Glu Tyr Tyr Ile Ala Asp
                 340             345             350

     GCC TCG GAG GAC CAG GTG TTT GTG TGT GTC AGT CAC AGC AAC AAC CGC    1284
     Ala Ser Glu Asp Gln Val Phe Val Cys Val Ser His Ser Asn Asn Arg
                 355             360             365

     ACC AAC CTC TAC ATC TCG GAG GCA GAG GGC TTG AAG TTC TCT CTG TCC    1332
     Thr Asn Leu Tyr Ile Ser Glu Ala Glu Gly Leu Lys Phe Ser Leu Ser
     370             375             380             385

     CTG GAG AAC GTG CTC TAC TAC ACC CCG GGA GGG GCC GGC AGT GAC ACC    1380
     Leu Glu Asn Val Leu Tyr Tyr Thr Pro Gly Gly Ala Gly Ser Asp Thr
                 390             395             400

     TTG GTG AGG TAC TTT GCA AAT GAA CCG TTT GCT GAC TTC CAT CGT GTG    1428
     Leu Val Arg Tyr Phe Ala Asn Glu Pro Phe Ala Asp Phe His Arg Val
                 405             410             415

     GAA GGG TTG CAG GGA GTC TAC ATT GCT ACT CTG ATT AAT GGT TCT ATG    1476
     Glu Gly Leu Gln Gly Val Tyr Ile Ala Thr Leu Ile Asn Gly Ser Met
                 420             425             430

     AAT GAG GAG AAC ATG AGA TCT GTC ATC ACC TTT GAC AAA GGG GGC ACC    1524
     Asn Glu Glu Asn Met Arg Ser Val Ile Thr Phe Asp Lys Gly Gly Thr
```

```
          435                 440                 445
TGG GAA TTT CTG CAG GCT CCA GCC TTC ACG GGG TAT GGA GAG AAA ATC    1572
Trp Glu Phe Leu Gln Ala Pro Ala Phe Thr Gly Tyr Gly Glu Lys Ile
   450                 455                 460                 465
AAC TGT GAG CTG TCC GAG GGC TGT TCC CTC CAC CTG GCC CAG CGC CTC    1620
Asn Cys Glu Leu Ser Glu Gly Cys Ser Leu His Leu Ala Gln Arg Leu
                    470                 475                 480
AGC CAG CTG CTC AAC CTC CAG CTC CGG AGG ATG CCC ATC CTG TCC AAG    1668
Ser Gln Leu Leu Asn Leu Gln Leu Arg Arg Met Pro Ile Leu Ser Lys
                   485                 490                 495
GAG TCG GCG CCT GGC CTC ATC ATT GCC ACG GGC TCA GTG GGA AAG AAC    1716
Glu Ser Ala Pro Gly Leu Ile Ile Ala Thr Gly Ser Val Gly Lys Asn
                   500                 505                 510
TTG GCT AGC AAG ACA AAC GTG TAC ATC TCT AGC AGT GCT GGA GCC AGG    1764
Leu Ala Ser Lys Thr Asn Val Tyr Ile Ser Ser Ser Ala Gly Ala Arg
             515                 520                 525
TGG CGA GAG GCA CTT CCT GGA CCT CAC TAC TAT ACA TGG GGA GAC CAT    1812
Trp Arg Glu Ala Leu Pro Gly Pro His Tyr Tyr Thr Trp Gly Asp His
530                 535                 540                 545
GGC GGC ATC ATC ATG GCC ATT GCC CAA GGC ATG GAA ACC AAC GAA CTG    1860
Gly Gly Ile Ile Met Ala Ile Ala Gln Gly Met Glu Thr Asn Glu Leu
                   550                 555                 560
AAG TAC AGT ACC AAC GAA GGG GAG ACC TGG AAA GCC TTC ACC TTC TCT    1908
Lys Tyr Ser Thr Asn Glu Gly Glu Thr Trp Lys Ala Phe Thr Phe Ser
             565                 570                 575
GAG AAG CCC GTG TTT GTG TAT GGG CTC CTC ACG GAA CCC GGC GAG AAG    1956
Glu Lys Pro Val Phe Val Tyr Gly Leu Leu Thr Glu Pro Gly Glu Lys
       580                 585                 590
```

AGC ACG GTC TTC ACC ATC TTT GGC TCC AAC AAG GAG AAC GTG CAC AGC   2004
Ser Thr Val Phe Thr Ile Phe Gly Ser Asn Lys Glu Asn Val His Ser
    595   600   605

TGG CTC ATC CTC CAG GTC AAT GCC ACA GAC GCC CTG GGG GTT CCT TGC   2052
Trp Leu Ile Leu Gln Val Asn Ala Thr Asp Ala Leu Gly Val Pro Cys
610   615   620   625

ACA GAG AAC GAC TAC AAG CTC TGG TCA CCA TCT GAT GAG CGG GGG AAT   2100
Thr Glu Asn Asp Tyr Lys Leu Trp Ser Pro Ser Asp Glu Arg Gly Asn
       630   635   640

GAG TGT TTG CTT GGA CAC AAG ACT GTT TTC AAA CGG AGG ACC CCG CAC   2148
Glu Cys Leu Leu Gly His Lys Thr Val Phe Lys Arg Arg Thr Pro His
       645   650   655

GCC ACA TGC TTT AAC GGA GAA GAC TTT GAC AGG CCG GTG GTT GTG TCC   2196
Ala Thr Cys Phe Asn Gly Glu Asp Phe Asp Arg Pro Val Val Val Ser
      660   665   670

AAC TGC TCC TGC ACC CGG GAG GAC TAT GAG TGT GAC TTT GGC TTC CGG   2244
Asn Cys Ser Cys Thr Arg Glu Asp Tyr Glu Cys Asp Phe Gly Phe Arg
    675   680   685

ATG AGT GAA GAC TTG GCA TTA GAG GTG TGT GTT CCA GAT CCA GGA TTT   2292
Met Ser Glu Asp Leu Ala Leu Glu Val Cys Val Pro Asp Pro Gly Phe
690   695   700   705

TCT GGA AAG TCC TCC CCT CCA GTG CCT TGT CCC GTG GGC TCT ACG TAC   2340
Ser Gly Lys Ser Ser Pro Pro Val Pro Cys Pro Val Gly Ser Thr Tyr
       710   715   720

AGG CGA TCA AGA GGC TAC CGG AAG ATT TCT GGG GAC ACC TGT AGT GGA   2388
Arg Arg Ser Arg Gly Tyr Arg Lys Ile Ser Gly Asp Thr Cys Ser Gly
       725   730   735

GGA GAT GTT GAG GCA CGG CTA GAA GGA GAG CTG GTC CCC TGT CCC CTG   2436

```
        Gly Asp Val Glu Ala Arg Leu Glu Gly Glu Leu Val Pro Cys Pro Leu
             740                 745                 750

        GCA GAA GAG AAC GAG TTC ATC CTG TAC GCC ACG CGC AAG TCC ATC CAC    2484
        Ala Glu Glu Asn Glu Phe Ile Leu Tyr Ala Thr Arg Lys Ser Ile His
             755                 760                 765

        CGC TAT GAC CTG GCT TCC GGA ACC ACG GAG CAG TTG CCC CTC ACT GGG    2532
        Arg Tyr Asp Leu Ala Ser Gly Thr Thr Glu Gln Leu Pro Leu Thr Gly
        770                 775                 780                 785

        TTC CGG GCA GCA GTG GCC CTG GAC TTT GAC TAT GAG CAC AAC TGC CTG    2580
        Leu Arg Ala Ala Val Ala Leu Asp Phe Asp Tyr Glu His Asn Cys Leu
                         790                 795                 800

        TAT TGG TCT GAC CTG GCC TTG GAC GTC ATC CAG CGC CTC TGT TTG AAC    2628
        Tyr Trp Ser Asp Leu Ala Leu Asp Val Ile Gln Arg Leu Cys Leu Asn
                     805                 810                 815

        GGG AGT ACA GGA CAA GAG GTG ATC ATC AAC TCT GAC CTG GAG ACG GTA    2676
        Gly Ser Thr Gly Gln Glu Val Ile Ile Asn Ser Asp Leu Glu Thr Val
             820                 825                 830

        GAA GCT TTG GCT TTT GAA CCC CTC AGC CAA TTA CTT TAC TGG GTG GAC    2724
        Glu Ala Leu Ala Phe Glu Pro Leu Ser Gln Leu Leu Tyr Trp Val Asp
             835                 840                 845

        GCA GGC TTT AAA AAG ATC GAG GTA GCC AAT CCA GAT GGT GAC TTC CGA    2772
        Ala Gly Phe Lys Lys Ile Glu Val Ala Asn Pro Asp Gly Asp Phe Arg
        850                 855                 860                 865

        CTC ACC GTC GTC AAT TCC TCG GTG CTG GAT CGG CCC CGG GCC CTG GTC    2820
        Leu Thr Val Val Asn Ser Ser Val Leu Asp Arg Pro Arg Ala Leu Val
                         870                 875                 880

        CTT GTG CCC CAA GAA GGG ATC ATG TTC TGG ACC GAC TGG GGA GAC CTG    2868
        Leu Val Pro Gln Glu Gly Ile Met Phe Trp Thr Asp Trp Gly Asp Leu
```

885   890   895

AAG CCT GGG ATT TAT CGG AGC AAC ATG GAC GGA TCT GCC GCC TAT CGC 2916
Lys Pro Gly Ile Tyr Arg Ser Asn Met Asp Gly Ser Ala Ala Tyr Arg

  900   905   910

CTC GTG TCG GAG GAT GTG AAG TGG CCC AAT GGC ATT TCC GTG GAC GAT 2964
Leu Val Ser Glu Asp Val Lys Trp Pro Asn Gly Ile Ser Val Asp Asp

  915   920   925

CAG TGG ATC TAC TGG ACG GAT GCC TAC CTG GAC TGC ATT GAG CGC ATC 3012
Gln Trp Ile Tyr Trp Thr Asp Ala Tyr Leu Asp Cys Ile Glu Arg Ile

930   935   940   945

ACG TTC AGC GGC CAG CAG CGC TCC GTC ATC CTG GAC AGA CTC CCG CAC 3060
Thr Phe Ser Gly Gln Gln Arg Ser Val Ile Leu Asp Arg Leu Pro His

    950   955   960

CCC TAT GCC ATT GCT GTC TTT AAG AAT GAG ATT TAC TGG GAT GAC TGG 3108
Pro Tyr Ala Ile Ala Val Phe Lys Asn Glu Ile Tyr Trp Asp Asp Trp

    965   970   975

TCA CAG CTC AGC ATA TTC CGA GCT TCT AAG TAC AGC GGG TCC CAG ATG 3156
Ser Gln Leu Ser Ile Phe Arg Ala Ser Lys Tyr Ser Gly Ser Gln Met

  980   985   990

GAG ATT CTG GCC AGC CAG CTC ACG GGG CTG ATG GAC ATG AAG ATC TTC 3204
Glu Ile Leu Ala Ser Gln Leu Thr Gly Leu Met Asp Met Lys Ile Phe

  995   1000   1005

TAC AAG GGG AAG AAC ACA GGA AGC AAT GCG TGT GTA CCC AGG CCG TGC 3252
Tyr Lys Gly Lys Asn Thr Gly Ser Asn Ala Cys Val Pro Arg Pro Cys

1010   1015   1020   1025

AGC CTG CTG TGC CTG CCC AGA GCC AAC AAC AGC AAA AGC TGC AGG TGT 3300
Ser Leu Leu Cys Leu Pro Arg Ala Asn Asn Ser Lys Ser Cys Arg Cys

    1030   1035   1040

CCA GAT GGC GTG GCC AGC AGT GTC CTC CCT TCC GGG GAC CTG ATG TGT   3348
Pro Asp Gly Val Ala Ser Ser Val Leu Pro Ser Gly Asp Leu Met Cys
           1045   1050   1055

GAC TGC CCT AAG GGC TAC GAG CTG AAG AAC AAC ACG TGT GTC AAA GAA   3396
Asp Cys Pro Lys Gly Tyr Glu Leu Lys Asn Asn Thr Cys Val Lys Glu
           1060   1065   1070

GAA GAC ACC TGT CTG CGC AAC CAG TAC CGC TGC AGC AAC GGG AAC TGC   3444
Glu Asp Thr Cys Leu Arg Asn Gln Tyr Arg Cys Ser Asn Gly Asn Cys
          1075   1080   1085

ATC AAC AGC ATC TGG TGG TGC GAT TTC GAC AAC GAC TGC GGA GAC ATG   3492
Ile Asn Ser Ile Trp Trp Cys Asp Phe Asp Asn Asp Cys Gly Asp Met
1090   1095   1100   1105

AGC GAC GAG AAG AAC TGC CCT ACC ACC ATC TGC GAC CTG GAC ACC CAG   3540
Ser Asp Glu Lys Asn Cys Pro Thr Thr Ile Cys Asp Leu Asp Thr Gln
           1110   1115   1120

TTC CGT TGC CAG GAG TCT GGG ACG TGC ATC CCG CTC TCC TAC AAA TGT   3588
Phe Arg Cys Gln Glu Ser Gly Thr Cys Ile Pro Leu Ser Tyr Lys Cys
          1125   1130   1135

GAC CTC GAG GAT GAC TGT GGG GAC AAC AGT GAC GAA AGG CAC TGT GAA   3636
Asp Leu Glu Asp Asp Cys Gly Asp Asn Ser Asp Glu Arg His Cys Glu
          1140   1145   1150

ATG CAC CAG TGC CGG AGC GAC GAA TAC AAC TGC AGC TCG GGC ATG TGC   3684
Met His Gln Cys Arg Ser Asp Glu Tyr Asn Cys Ser Ser Gly Met Cys
          1155   1160   1165

ATC CGC TCC TCC TGG GTG TGC GAC GGG GAC AAC GAC TGC AGG GAC TGG   3732
Ile Arg Ser Ser Trp Val Cys Asp Gly Asp Asn Asp Cys Arg Asp Trp
1170   1175   1180   1185

TCC GAC GAG GCC AAC TGC ACA GCC ATC TAT CAC ACC TGT GAG GCC TCC   3780

Ser Asp Glu Ala Asn Cys Thr Ala Ile Tyr His Thr Cys Glu Ala Ser

           1190   1195   1200

AAC TTC CAG TGC CGC AAC GGG CAC TGC ATC CCC CAG CGG TGG GCG TGT   3828

Asn Phe Gln Cys Arg Asn Gly His Cys Ile Pro Gln Arg Trp Ala Cys

           1205   1210   1215

GAC GGC GAC GCC GAC TGC CAG GAT GGC TCT GAT GAG GAT CCA GCC AAC   3876

Asp Gly Asp Ala Asp Cys Gln Asp Gly Ser Asp Glu Asp Pro Ala Asn

         1220   1225   1230

TGT GAG AAG AAG TGC AAC GGC TTC CGC TGC CCG AAC GGC ACC TGC ATT   3924

Cys Glu Lys Lys Cys Asn Gly Phe Arg Cys Pro Asn Gly Thr Cys Ile

      1235   1240   1245

CCC TCC ACC AAG CAC TGT GAC GGC CTG CAC GAT TGC TCG GAC GGC TCC   3972

Pro Ser Thr Lys His Cys Asp Gly Leu His Asp Cys Ser Asp Gly Ser

1250   1255   1260   1265

GAC GAG CAG CAC TGC GAG CCC CTG TGT ACA CGG TTC ATG GAC TTC GTG   4020

Asp Glu Gln His Cys Glu Pro Leu Cys Thr Arg Phe Met Asp Phe Val

           1270   1275   1280

TGT AAG AAC CGC CAG CAG TGC CTC TTC CAC TCC ATG GTG TGC GAT GGG   4068

Cys Lys Asn Arg Gln Gln Cys Leu Phe His Ser Met Val Cys Asp Gly

          1285   1290   1295

ATC ATC CAG TGC CGT GAC GGC TCC GAC GAG GAC CCA GCC TTT GCA GGA   4116

Ile Ile Gln Cys Arg Asp Gly Ser Asp Glu Asp Pro Ala Phe Ala Gly

        1300   1305   1310

TGC TCC CGA GAC CCC GAG TTC CAC AAG GTG TGC GAT GAG TTC GGC TTC   4164

Cys Ser Arg Asp Pro Glu Phe His Lys Val Cys Asp Glu Phe Gly Phe

      1315   1320   1325

CAG TGT CAG AAC GGC GTG TGC ATC AGC TTG ATC TGG AAG TGC GAC GGG   4212

Gln Cys Gln Asn Gly Val Cys Ile Ser Leu Ile Trp Lys Cys Asp Gly

```
      1330              1335              1340              1345
ATG GAT GAC TGC GGG GAC TAC TCC GAC GAG GCC AAC TGT GAA AAC CCC    4260
Met Asp Asp Cys Gly Asp Tyr Ser Asp Glu Ala Asn Cys Glu Asn Pro
                  1350              1355              1360
ACA GAA GCC CCC AAC TGC TCC CGC TAC TTC CAG TTC CGG TGT GAC AAT    4308
Thr Glu Ala Pro Asn Cys Ser Arg Tyr Phe Gln Phe Arg Cys Asp Asn
              1365              1370              1375
GGC CAC TGC ATC CCC AAC AGG TGG AAG TGT GAC AGG GAG AAT GAC TGT    4356
Gly His Cys Ile Pro Asn Arg Trp Lys Cys Asp Arg Glu Asn Asp Cys
                  1380              1385              1390
GGG GAC TGG TCC GAC GAG AAG GAC TGT GGA GAT TCA CAT GTA CTT CCG    4404
Gly Asp Trp Ser Asp Glu Lys Asp Cys Gly Asp Ser His Val Leu Pro
              1395              1400              1405
TCT ACG ACT CCT GCA CCC TCC ACG TGT CTG CCC AAT TAC TAC CGC TGC    4452
Ser Thr Thr Pro Ala Pro Ser Thr Cys Leu Pro Asn Tyr Tyr Arg Cys
      1410              1415              1420              1425
GGC GGG GGG GCC TGC GTG ATA GAC ACG TGG GTT TGT GAC GGG TAC CGA    4500
Gly Gly Gly Ala Cys Val Ile Asp Thr Trp Val Cys Asp Gly Tyr Arg
                  1430              1435              1440
GAT TGC GCA GAT GGA TCC GAC GAG GAA GCC TGC CCC TCG CTC CCC AAT    4548
Asp Cys Ala Asp Gly Ser Asp Glu Glu Ala Cys Pro Ser Leu Pro Asn
              1445              1450              1455
GTC ACT GCC ACC TCC TCC CCC TCC CAG CCT GGA CGA TGC GAC CGA TTT    4596
Val Thr Ala Thr Ser Ser Pro Ser Gln Pro Gly Arg Cys Asp Arg Phe
              1460              1465              1470
GAG TTT GAG TGC CAC CAG CCA AAG AAG TGC ATC CCT AAC TGG AGA CGC    4644
Glu Phe Glu Cys His Gln Pro Lys Lys Cys Ile Pro Asn Trp Arg Arg
      1775              1480              1485
```

TGT GAC GGC CAT CAG GAT TGC CAG GAT GGC CAG GAC GAG GCC AAC TGC   4692
Cys Asp Gly His Gln Asp Cys Gln Asp Gly Gln Asp Glu Ala Asn Cys
      1490   1495   1500   1505

CCC ACT CAC AGC ACC TTG ACC TGC ATG AGC TGG GAG TTC AAG TGT GAG   4740
Pro Thr His Ser Thr Leu Thr Cys Met Ser Trp Glu Phe Lys Cys Glu
            1510   1515   1520

GAT GGC GAG GCC TGC ATC GTG CTG TCA GAA CGC TGC GAC GGC TTC CTG   4788
Asp Gly Glu Ala Cys Ile Val Leu Ser Glu Arg Cys Asp Gly Phe Leu
        1525   1530   1535

GAC TGC TCA GAT GAG AGC GAC GAG AAG GCC TGC AGT GAT GAG TTA ACT   4836
Asp Cys Ser Asp Glu Ser Asp Glu Lys Ala Cys Ser Asp Glu Leu Thr
      1540   1545   1550

GTA TAC AAA GTA CAG AAT CTT CAG TGG ACA GCT GAC TTC TCT GGG AAT   4884
Val Tyr Lys Val Gln Asn Leu Gln Trp Thr Ala Asp Phe Ser Gly Asn
    1555   1560   1565

GTC ACT TTG ACC TGG ATG CGG CCC AAA AAA ATG CCC TCT GCT GCT TGT   4932
Val Thr Leu Thr Trp Met Arg Pro Lys Lys Met Pro Ser Ala Ala Cys
      1570   1575   1580   1585

GTA TAC AAC GTG TAC TAT AGA GTT GTT GGA GAG AGC ATA TGG AAG ACT   4980
Val Tyr Asn Val Tyr Tyr Arg Val Val Gly Glu Ser Ile Trp Lys Thr
        1590   1595   1600

CTG GAG ACT CAC AGC AAT AAG ACA AAC ACT GTA TTA AAA GTG TTC AAA   5028
Leu Glu Thr His Ser Asn Lys Thr Asn Thr Val Leu Lys Val Leu Lys
        1605   1610   1615

CCA GAT ACC ACC TAC CAG GTT AAA GTG CAG GTT CAG TGC CTG AGC AAG   5076
Pro Asp Thr Thr Tyr Gln Val Lys Val Gln Val Gln Cys Leu Ser Lys
        1620   1625   1630

GTG CAC AAC ACC AAT GAC TTT GTG ACC TTG AGA ACT CCA GAG GGA TTG   5124

Val His Asn Thr Asn Asp Phe Val Thr Leu Arg Thr Pro Glu Gly Leu

   1635   1640   1645

CCA GAC GCC CCT CAG AAC CTC CAG CTG TCG CTC CAC GGG GAA GAG GAA   5172

Pro Asp Ala Pro Gln Asn Leu Gln Leu Ser Leu His Gly Glu Glu Glu

1650   1655   1660   1665

GGT GTG ATT GTG GGC CAC TGG AGC CCT CCC ACC CAC ACC CAC GGC CTC   5220

Gly Val Ile Val Gly His Trp Ser Pro Pro Thr His Thr His Gly Leu

        1670   1675   1680

ATT CGC GAA TAC ATT GTA GAG TAT AGC AGG AGT GGT TCC AAG GTG TGG   5268

Ile Arg Glu Tyr Ile Val Glu Tyr Ser Arg Ser Gly Ser Lys Val Trp

        1685   1690   1695

ACT TCA GAA AGG GCT GCT AGT AAC TTT ACA GAA ATA AAG AAC TTG TTG   5316

Thr Ser Glu Arg Ala Ala Ser Asn Phe Thr Glu Ile Lys Asn Leu Leu

   1700   1705   1710

GTC AAC ACC CTG TAC ACC GTC AGA GTG GCT GCG GTG ACG AGT CGT GGG   5364

Val Asn Thr Leu Tyr Thr Val Arg Val Ala Ala Val Thr Ser Arg Gly

   1715   1720   1725

ATA GGA AAC TGG AGC GAT TCC AAA TCC ATT ACC ACC GTG AAA GGA AAA   5412

Ile Gly Asn Trp Ser Asp Ser Lys Ser Ile Thr Thr Val Lys Gly Lys

1730   1735   1740   1745

GCG ATC CCG CCA CCA AAT ATC CAC ATT GAC AAC TAC GAT GAA AAT TCC   5460

Ala Ile Pro Pro Pro Asn Ile His Ile Asp Asn Tyr Asp Glu Asn Ser

        1750   1755   1760

CTG AGT TTT ACC CTG ACC GTG GAT GGG AAC ATC AAG GTG AAT GGC TAT   5508

Leu Ser Phe Thr Leu Thr Val Asp Gly Asn Ile Lys Val Asn Gly Tyr

        1765   1770   1775

GTG GTG AAC CTT TTC TGG GCA TTT GAC ACC CAC AAA CAA GAG AAG AAA   5556

Val Val Asn Leu Phe Trp Ala Phe Asp Thr His Lys Gln Glu Lys Lys

```
          1780                1785                1790
ACC ATG AAC TTC CAA GGG AGC TCA GTG TCC CAC AAA GTT GGC AAT CTG    5604
Thr Met Asn Phe Gln Gly Ser Ser Val Ser His Lys Val Gly Asn Leu
       1795                1800                1805
ACA GCA CAG ACG GCC TAT GAG ATT TCC GCC TGG GCC AAG ACT GAC TTG    5652
Thr Ala Gln Thr Ala Tyr Glu Ile Ser Ala Trp Ala Lys Thr Asp Leu
  1810                1815                1820                1825
GGC GAT AGT CCT CTG TCA TTT GAG CAT GTC ACG ACC AGA GGG GTT CGC    5700
Gly Asp Ser Pro Leu Ser Phe Glu His Val Thr Thr Arg Gly Val Arg
                1830                1835                1840
CCA CCT GCT CCT AGC CTC AAG GCC AGG GCT ATC AAT CAG ACT GCA GTG    5748
Pro Pro Ala Pro Ser Leu Lys Ala Arg Ala Ile Asn Gln Thr Ala Val
             1845                1850                1855
GAA TGC ACC TGG ACA GGC CCC AGG AAT GTG GTG TAT GGC ATT TTC TAT    5796
Glu Cys Thr Trp Thr Gly Pro Arg Asn Val Val Tyr Gly Ile Phe Tyr
          1860                1865                1870
GCC ACA TCC TTC CTG GAC CTC TAC CGC AAC CCA AGC AGC CTG ACC ACG    5844
Ala Thr Ser Phe Leu Asp Leu Tyr Arg Asn Pro Ser Ser Leu Thr Thr
       1875                1880                1885
CCG CTG CAC AAC GCA ACC GTG CTC GTC GGT AAG GAT GAG CAG TAT CTG    5892
Pro Leu His Asn Ala Thr Val Leu Val Gly Lys Asp Glu Gln Tyr Leu
  1890                1895                1900                1905
TTT CTG GTC CGG GTG GTG ATG CCC TAC CAA GGG CCG TCC TCG GAC TAC    5940
Phe Leu Val Arg Val Val Met Pro Tyr Gln Gly Pro Ser Ser Asp Tyr
                1910                1915                1920
GTG GTC GTG AAG ATG ATC CCG GAC AGC AGG CTT CCT CCC CGG CAC CTG    5988
Val Val Val Lys Met Ile Pro Asp Ser Arg Leu Pro Pro Arg His Leu
             1925                1930                1935
```

```
CAT GCC GTT CAC ACC GGC AAG ACC TCG GCC GTC ATC AAG TGG GAG TCG    6036
His Ala Val His Thr Gly Lys Thr Ser Ala Val Ile Lys Trp Glu Ser
         1940            1945            1950

CCC TAC GAC TCT CCT GAC CAG GAC CTG TTC TAT GCG ATC GCA GTT AAA    6084
Pro Tyr Asp Ser Pro Asp Gln Asp Leu Phe Tyr Ala Ile Ala Val Lys
         1955            1960            1965

GAT CTG ATA CGA AAG ACG GAC CGG AGC TAC AAA GTC AAG TCC CGC AAC    6132
Asp Leu Ile Arg Lys Thr Asp Arg Ser Tyr Lys Val Lys Ser Arg Asn
1970            1975            1980            1985

AGC ACC GTG GAG TAC ACC CTG AGC AAG CTG GAG CCC GGA GGG AAA TAC    6180
Ser Thr Val Glu Tyr Thr Leu Ser Lys Leu Glu Pro Gly Gly Lys Tyr
                1990            1995            2000

CAC GTC ATT GTG CAG CTG GGG AAC ATG AGC AAA GAT GCC AGT GTG AAG    6228
His Val Ile Val Gln Leu Gly Asn Met Ser Lys Asp Ala Ser Val Lys
            2005            2010            2015

ATC ACC ACC GTT TCG TTA TCG GCA CCC GAT GCC TTA AAA ATC ATA ACA    6276
Ile Thr Thr Val Ser Leu Ser Ala Pro Asp Ala Leu Lys Ile Ile Thr
            2020            2025            2030

GAA AAT GAC CAC GTC CTT CTC TTC TGG AAA AGT CTA GCT CTA AAG GAA    6324
Glu Asn Asp His Val Leu Leu Phe Trp Lys Ser Leu Ala Leu Lys Glu
         2035            2040            2045

AAG TAT TTT AAC GAA AGC AGG GGC TAC GAG ATA CAC ATG TTT GAT AGC    6372
Lys Tyr Phe Asn Glu Ser Arg Gly Tyr Glu Ile His Met Phe Asp Ser
2050            2055            2060            2065

GCC ATG AAT ATC ACC GCA TAC CTT GGG AAT ACT ACT GAC AAT TTC TTT    6420
Ala Met Asn Ile Thr Ala Tyr Leu Gly Asn Thr Thr Asp Asn Phe Phe
                2070            2075            2080

AAA ATT TCC AAC CTG AAG ATG GGT CAC AAT TAC ACA TTC ACG GTC CAG    6468
```

```
            Lys Ile Ser Asn Leu Lys Met Gly His Asn Tyr Thr Phe Thr Val Gln
                         2085                2090                2095

            GCA CGA TGC CTT TTG GGC AGC CAG ATC TGC GGG GAG CCT GCC GTG CTA     6516
            Ala Arg Cys Leu Leu Gly Ser Gln Ile Cys Gly Glu Pro Ala Val Leu
                         2100                2105                2110

            CTG TAT GAT GAG CTG GGG TCT GGT GGC GAT GCG TCG GCG ATG CAG GCT     6564
            Leu Tyr Asp Glu Leu Gly Ser Gly Gly Asp Ala Ser Ala Met Gln Ala
                       2115                2120                2125

            GCC AGG TCT ACT GAT GTC GCC GCC GTG GTG GTG CCC ATC CTG TTT CTG     6612
            Ala Arg Ser Thr Asp Val Ala Ala Val Val Val Pro Ile Leu Phe Leu
            2130                2135                2140                2145

            ATA CTG CTG AGC CTG GGG GTC GGG TTT GCC ATC CTG TAC ACG AAG CAT     6660
            Ile Leu Leu Ser Leu Gly Val Gly Phe Ala Ile Leu Tyr Thr Lys His
                         2150                2155                2160

            CGG AGG CTG CAG AGC AGC TTC ACC GCC TTC GCC AAC AGC CAC TAC AGC     6708
            Arg Arg Leu Gln Ser Ser Phe Thr Ala Phe Ala Asn Ser His Tyr Ser
                         2165                2170                2175

            TCC AGA CTC GGC TCC GCC ATC TTC TCC TCT GGG GAT GAC TTG GGG GAG     6756
            Ser Arg Leu Gly Ser Ala Ile Phe Ser Ser Gly Asp Asp Leu Gly Glu
                         2180                2185                2190

            GAT GAT GAA GAT GCT CCT ATG ATC ACT GGA TTT TCG GAC GAC GTC CCC     6804
            Asp Asp Glu Asp Ala Pro Met Ile Thr Gly Phe Ser Asp Asp Val Pro
                       2195                2200                2205

            ATG GTG ATA GCC TGAAAGAGCT TTCCTCACTA GAAACCAAAT GGTGTAAATA         6856
            Met Val Ile Ala
            2210

            TTTTATTTGA TAAAGATAGT TGATGGTTTA TTTTAAAAGA TGCACTTTGA GTTGCAATAT 6916
            GTTATTTTTA TATGGGCCAA AAACAAAAGC AAAAAAAAAA AAAAA               6961
```

Sequence ID No. 4

Length of the Sequence: 300

Type: nucleic acid

Strandedness: double

Topology: linear

Molecular type: cDNA to mRNA

Sequence:

```
ATATCCACAT TGACAGCTAT GGTGAAAATT ATCTAAGCTT CACCCTGACC ATGGAGAGTG   60

ATATCAAGGT GAATGGCTAT GTGGTGAACC TTTTCTGGGC ATTTGACACC CACAAGCAAG  120

AGAGGAGAAC TTTGAACTTC CGAGGAAGCA TATTGTCACA CAAAGTTGGC AATCTGACAG  180

CTCATACATC CTATGAGATT TCTGCCTGGG CCAAGACTGA CTTGGGGGAT AGCCCTCTGG  240

CATTTGAGCA TGTTATGACC AGAGGGGTTC GCCCACCTGC ACCTAGCCTC AAGGCCAAAG  300
```

Sequence ID No. 5

Length of the Sequence: 6642

Type: nucleic acid

Strandedness: double

Topology: linear

Molecular type: cDNA to mRNA

Sequence:

```
ATGGCGACAC GGAGCAGCAG GAGGGAGTCG CGACTCCCGT TCCTATTCAC CCTGGTCGCA   60

CTGCTGCCGC CCGGAGCTCT CTGCGAAGTC TGGACGCAGA GGCTGCACGG CGGCAGCGCG  120

CCCTTGCCCC AGGACCGGGG CTTCCTCGTG GTGCAGGGCG ACCCGCGCGA GCTGCGGCTG  180

TGGGCGCGCG GGGATGCCAG GGGGGCGAGC CGCGCGGACG AGAAGCCGCT CCGGAGGAAA  240

CGGAGCGCTG CCCTGCAGCC CGAGCCCATC AAGGTGTACG GACAGGTTAG TCTGAATGAT  300

TCCCACAATC AGATGGTGCT GCACTGGGCT GGAGAGAAAA GCAACGTGAT CGTGGCCTTG  360

GCCCGAGATA GCCTGGCATT GGCGAGGCCC AAGAGCAGTG ATGTGTACGT GTCTTACGAC  420

TATGGAAAAT CATTCAAGAA AATTTCAGAC AAGTTAAACT TTGGCTTGGG AAATAGGAGT  480
```

```
GAAGCTGTTA TCGCCCAGTT CTACCACAGC CCTGCGGACA ACAAGCGGTA CATCTTTGCA    540

GACGCTTATG CCCAGTACCT CTGGATCACG TTTGACTTCT GCAACACTCT TCAAGGCTTT    600

TCCATCCCAT TTCGGGCAGC TGATCTCCTC CTACACAGTA AGGCCTCCAA CCTTCTCTTG    660

GGCTTTGACA GGTCCCACCC CAACAAGCAG CTGTGGAAGT CAGATGACTT TGGCCAGACC    720

TGGATCATGA TTCAGGAACA TGTCAAGTCC TTTTCTTGGG GAATTGATCC CTATGACAAA    780

CCAAATACCA TCTACATTGA ACGACACGAA CCCTCTGGCT ACTCCACTGT CTTCCGAAGT    840

ACAGATTTCT TCCAGTCCCG GGAAAACCAG GAAGTGATCC TTGAGGAAGT GAGAGATTTT    900

CAGCTTCGGG ACAAGTACAT GTTTGCTACA AAGGTGGTGC ATCTCTTGGG CAGTGAACAG    960

CAGTCTTCTG TCCAGCTCTG GGTCTCCTTT GGCCGGAAGC CCATGAGAGC AGCCCAGTTT    1020

GTCACAAGAC ATCCTATTAA TGAATATTAC ATCGCAGATG CCTCCGAGGA CCAGGTGTTT    1080

GTGTGTGTCA GCCACAGTAA CAACCGCACC AATTTATACA TCTCAGAGGC AGAGGGGCTG    1140

AAGTTCTCCC TGTCCTTGGA GAACGTGCTC TATTACAGCC CAGGAGGGGC CGGCAGTGAC    1200

ACCTTGGTGA GGTATTTTGC AAATGAACCA TTTGCTGACT TCCACCGAGT GGAAGGATTG    1260

CAAGGAGTCT ACATTGCTAC TCTGATTAAT GGTTCTATGA ATGAGGAGAA CATGAGATCG    1320

GTCATCACCT TTGACAAAGG GGGAACCTGG GAGTTTCTTC AGGCTCCAGC CTTCACGGGA    1380

TATGGAGAGA AAATCAATTG TGAGCTTTCC CAGGGCTGTT CCCTTCATCT GGCTCAGCGC    1440

CTCAGTCAGC TCCTCAACCT CCAGCTCCGG AGAATGCCCA TCCTGTCCAA GGAGTCGGCT    1500

CCAGGCCTCA TCATCGCCAC TGGCTCAGTG GGAAAGAACT TGGCTAGCAA GACAAACGTG    1560

TACATCTCTA GCAGTGCTGG AGCCAGGTGG CGAGAGGCAC TTCCTGGACC TCACTACTAC    1620

ACATGGGGAG ACCACGGCGG AATCATCACG GCCATTGCCC AGGGCATGGA AACCAACGAG    1680

CTAAAATACA GTACCAATGA AGGGGAGACC TGGAAAACAT TCATCTTCTC TGAGAAGCCA    1740

GTGTTTGTGT ATGGCCTCCT CACAGAACCT GGGGAGAAGA GCACTGTCTT CACCATCTTT    1800

GGCTCGAACA AAGAGAATGT CCACAGCTGG CTGATCCTCC AGGTCAATGC CACGGATGCC    1860

TTGGGAGTTC CCTGCACAGA GAATGACTAC AAGCTGTGGT CACCATCTGA TGAGCGGGGG    1920

AATGAGTGTT TGCTGGGACA CAAGACTGTT TTCAAACGGC GGACCCCCCA TGCCACATGC    1980

TTCAATGGAG AGGACTTTGA CAGGCCGGTG GTCGTGTCCA ACTGCTCCTG CACCCGGGAG    2040

GACTATGAGT GTGACTTCGG TTTCAAGATG AGTGAAGATT TGTCATTAGA GGTTTGTGTT    2100

CCAGATCCGG AATTTTCTGG AAAGTCATAC TCCCCTCCTG TGCCTTGCCC TGTGGGTTCT    2160
```

ACTTACAGGA GAACGAGAGG CTACCGGAAG ATTTCTGGGG ACACTTGTAG CGGAGGAGAT 2220

GTTGAAGCGC GACTGGAAGG AGAGCTGGTC CCCTGTCCCC TGGCAGAAGA GAACGAGTTC 2280

ATTCTGTATG CTGTGAGGAA ATCCATCTAC CGCTATGACC TGGCCTCGGG AGCCACCGAG 2340

CAGTTGCCTC TCACCGGGCT ACGGGCAGCA GTGGCCCTGG ACTTTGACTA TGAGCACAAC 2400

TGTTTGTATT GGTCCGACCT GGCCTTGGAC GTCATCCAGC GCCTCTGTTT GAATGGAAGC 2460

ACAGGGCAAG AGGTGATCAT CAATTCTGGC CTGGAGACAG TAGAAGCTTT GGCTTTTGAA 2520

CCCCTCAGCC AGCTGCTTTA CTGGGTAGAT GCAGGCTTCA AAAAGATTGA GGTAGCTAAT 2580

CCAGATGGCG ACTTCCGACT CACAATCGTC AATTCCTCTG TGCTTGATCG TCCCAGGGCT 2640

CTGGTCCTCG TGCCCCAAGA GGGGGTGATG TTCTGGACAG ACTGGGGAGA CCTGAAGCCT 2700

GGGATTTATC GGAGCAATAT GGATGGTTCT GCTGCCTATC ACCTGGTGTC TGAGGATGTG 2760

AAGTGGCCCA ATGGCATCTC TGTGGACGAC CAGTGGATTT ACTGGACGGA TGCCTACCTG 2820

GAGTGCATAG AGCGGATCAC GTTCAGTGGC CAGCAGCGCT CTGTCATTCT GGACAACCTC 2880

CCGCACCCCT ATGCCATTGC TGTCTTTAAG AATGAAATCT ACTGGGATGA CTGGTCACAG 2940

CTCAGCATAT TCCGAGCTTC CAAATACAGT GGGTCCCAGA TGGAGATTCT GGCAAACCAG 3000

CTCACGGGGC TCATGGACAT GAAGATTTTC TACAAGGGGA AGAACACTGG AAGCAATGCC 3060

TGTGTGCCCA GGCCATGCAG CCTGCTGTGC CTGCCCAAGG CCAACAACAG TAGAAGCTGC 3120

AGGTGTCCAG AGGATGTGTC CAGCAGTGTG CTTCCATCAG GGACCTGAT GTGTGACTGC 3180

CCTCAGGGCT ATCAGCTCAA GAACAATACC TGTGTCAAAG AAGAGAACAC CTGTCTTCGC 3240

AACCAGTATC GCTGCAGCAA CGGGAACTGT ATCAACAGCA TTTGGTGGTG TGACTTTGAC 3300

AACGACTGTG GAGACATGAG CGATGAGAGA AACTGCCCTA CCACCATCTG TGACCTGGAC 3360

ACCCAGTTTC GTTGCCAGGA GTCTGGGACT TGTATCCCAC TGTCCTATAA ATGTGACCTT 3420

GAGGATGACT GTGGAGACAA CAGTGATGAA AGTCATTGTG AAATGCACCA GTGCCGGAGT 3480

GACGAGTACA ACTGCAGTTC CGGCATGTGC ATCCGCTCCT CCTGGGTATG TGACGGGGAC 3540

AACGACTGCA GGGACTGGTC TGATGAAGCC AACTGTACCG CCATCTATCA CACCTGTGAG 3600

GCCTCCAACT TCCAGTGCCG AAACGGGCAC TGCATCCCCC AGCGGTGGGC GTGTGACGGG 3660

GATACGGACT GCCAGGATGG TTCCGATGAG GATCCAGTCA ACTGTGAGAA GAAGTGCAAT 3720

GGATTCCGCT GCCCAAACGG CACTTGCATC CCATCCAGCA AACATTGTGA TGGTCTGCGT 3780

GATTGCTCTG ATGGCTCCGA TGAACAGCAC TGCGAGCCCC TCTGTACGCA CTTCATGGAC 3840

```
TTTGTGTGTA AGAACCGCCA GCAGTGCCTG TTCCACTCCA TGGTCTGTGA CGGAATCATC   3900
CAGTGCCGCG ACGGGTCCGA TGAGGATGCG GCGTTTGCAG GATGCTCCCA AGATCCTGAG   3960
TTCCACAAGG TATGTGATGA GTTCGGTTTC CAGTGTCAGA ATGGAGTGTG CATCAGTTTG   4020
ATTTGGAAGT GCGACGGGAT GGATGATTGC GGCGATTATT CTGATGAAGC CAACTGCGAA   4080
AACCCCACAG AAGCCCCAAA CTGCTCCCGC TACTTCCAGT TTCGGTGTGA GAATGGCCAC   4140
TGCATCCCCA ACAGATGGAA ATGTGACAGG GAGAACGACT GTGGGGACTG GTCTGATGAG   4200
AAGGATTGTG GAGATTCACA TATTCTTCCC TTCTCGACTC CTGGGCCCTC CACGTGTCTG   4260
CCCAATTACT ACCGCTGCAG CAGTGGGACC TGCGTGATGG ACACCTGGGT GTGCGACGGG   4320
TACCGAGATT GTGCAGATGG CTCTGACGAG GAAGCCTGCC CCTTGCTTGC AAACGTCACT   4380
GCTGCCTCCA CTCCCACCCA ACTTGGGCGA TGTGACCGAT TTGAGTTCGA ATGCCACCAA   4440
CCGAAGACGT GTATTCCCAA CTGGAAGCGC TGTGACGGCC ACCAAGATTG CCAGGATGGC   4500
CGGGACGAGG CCAATTGCCC CACACACAGC ACCTTGACTT GCATGAGCAG GGAGTTCCAG   4560
TGCGAGGACG GGGAGGCCTG CATTGTGCTC TCGGAGCGCT GCGACGGCTT CCTGGACTGC   4620
TCGGACGAGA GCGATGAAAA GGCCTGCAGT GATGAGTTGA CTGTGTACAA AGTACAGAAT   4680
CTTCAGTGGA CAGCTGACTT CTCTGGGGAT GTGACTTTGA CCTGGATGAG GCCCAAAAAA   4740
ATGCCCTCTG CATCTTGTGT ATATAATGTC TACTACAGGG TGGTTGGAGA GAGCATATGG   4800
AAGACTCTGG AGACCCACAG CAATAAGACA AACACTGTAT TAAAAGTCTT GAAACCAGAT   4860
ACCACGTATC AGGTTAAAGT ACAGGTTCAG TGTCTCAGCA AGGCACACAA CACCAATGAC   4920
TTTGTGACCC TGAGGACCCC AGAGGGATTG CCAGATGCCC CTCGAAATCT CCAGCTGTCA   4980
CTCCCCAGGG AAGCAGAAGG TGTGATTGTA GGCCACTGGG CTCCTCCCAT CCACACCCAT   5040
GGCCTCATCC GTGAGTACAT TGTAGAATAC AGCAGGAGTG GTTCCAAGAT GTGGGCCTCC   5100
CAGAGGGCTG CTAGTAACTT TACAGAAATC AAGAACTTAT TGGTCAACAC TCTATACACC   5160
GTCAGAGTGG CTGCGGTGAC TAGTCGTGGA ATAGGAAACT GGAGCGATTC TAAATCCATT   5220
ACCACCATAA AAGGAAAAGT GATCCCACCA CCAGATATCC ACATTGACAG CTATGGTGAA   5280
AATTATCTAA GCTTCACCCT GACCATGGAG AGTGATATCA AGGTGAATGG CTATGTGGTG   5340
AACCTTTTCT GGGCATTTGA CACCCACAAG CAAGAGAGGA GAACTTTGAA CTTCCGAGGA   5400
AGCATATTGT CACACAAAGT TGGCAATCTG ACAGCTCATA CATCCTATGA GATTTCTGCC   5460
TGGGCCAAGA CTGACTTGGG GGATAGCCCT CTGGCATTTG AGCATGTTAT GACCAGAGGG   5520
```

```
GTTCGCCCAC CTGCACCTAG CCTCAAGGCC AAAGCCATCA ACCAGACTGC AGTGGAATGT   5580

ACCTGGACCG GCCCCCGGAA TGTGGTTTAT GGTATTTTCT ATGCCACGTC CTTTCTTGAC   5640

CTCTATCGCA ACCCGAAGAG CTTGACTACT TCACTCCACA ACAAGACGGT CATTGTCAGT   5700

AAGGATGAGC AGTATTTGTT TCTGGTCCGT GTAGTGGTAC CCTACCAGGG GCCATCCTCT   5760

GACTACGTTG TAGTGAAGAT GATCCCGGAC AGCAGGCTTC CACCCCGTCA CCTGCATGTG   5820

GTTCATACGG GCAAAACCTC CGTGGTCATC AAGTGGGAAT CACCGTATGA CTCTCCTGAC   5880

CAGGACTTGT TGTATGCAAT TGCAGTCAAA GATCTCATAA GAAAGACTGA CAGGAGCTAC   5940

AAAGTAAAAT CCCGTAACAG CACTGTGGAA TACACCCTTA ACAAGTTGGA GCCTGGCGGG   6000

AAATACCACA TCATTGTCCA ACTGGGGAAC ATGAGCAAAG ATTCCAGCAT AAAAATTACC   6060

ACAGTTTCAT TATCAGCACC TGATGCCTTA AAAATCATAA CAGAAAATGA TCATGTTCTT   6120

CTCTTTTGGA AAAGCCTGGC TTTAAAGGAA AAGCATTTTA ATGAAAGCAG GGGCTATGAG   6180

ATACACATGT TTGATAGTGC CATGAATATC ACAGCTTACC TTGGGAATAC TACTGACAAT   6240

TTCTTTAAAA TTTCCAACCT GAAGATGGGT CATAATTACA CGTTCACCGT CCAAGCAAGA   6300

TGCCTTTTTG GCAACCAGAT CTGTGGGGAG CCTGCCATCC TGCTGTACGA TGAGCTGGGG   6360

TCTCGTGCAG ATGCATCTGC AACGCAGGCT GCCAGATCTA CGGATGTTGC TGCTGTGGTG   6420

GTGCCCATCT TATTCCTGAT ACTGCTGAGC CTGGGGGTGG GGTTTGCCAT CCTGTACACG   6480

AAGCACCGGA GGCTGCAGAG CAGCTTCACC GCCTTCGCCA ACAGCCACTA CAGCTCCAGG   6540

CTGGGGTCCG CAATCTTCTC CTCTGGGGAT GACCTGGGGG AAGATGATGA AGATGCCCCT   6600

ATGATAACTG GATTTTCAGA TGACGTCCCC ATGGTGATAG CC                      6642
```

Sequence ID No. 6

Length of the Sequence: 2214

Type: amino acid

Topology: linear

Molecular type: Protein

Sequence:

```
Met Ala Thr Arg Ser Ser Arg Arg Glu Ser Arg Leu Pro Phe Leu Phe
                     5               10                  15

Thr Leu Val Ala Leu Leu Pro Pro Gly Ala Leu Cys Glu Val Trp Thr
```

<pre>
              20   25   30
    Gln Arg Leu His Gly Gly Ser Ala Pro Leu Pro Gln Asp Arg Gly Phe
              35   40   45
    Leu Val Val Gln Gly Asp Pro Arg Glu Leu Arg Leu Trp Ala Arg Gly
              50   55   60
    Asp Ala Arg Gly Ala Ser Arg Ala Asp Glu Lys Pro Leu Arg Arg Lys
         65   70   75   80
    Arg Ser Ala Ala Leu Gln Pro Glu Pro Ile Lys Val Tyr Gly Gln Val
                        85   90   95
    Ser Leu Asn Asp Ser His Asn Gln Met Val Val His Trp Ala Gly Glu
                   100   105   110
    Lys Ser Asn Val Ile Val Ala Leu Ala Arg Asp Ser Leu Ala Leu Ala
                   115   120   125
    Arg Pro Lys Ser Ser Asp Val Tyr Val Ser Tyr Asp Tyr Gly Lys Ser
              130   135   140
    Phe Lys Lys Ile Ser Asp Lys Leu Asn Phe Gly Leu Gly Asn Arg Ser
         145   150   155   160
    Glu Ala Val Ile Ala Gln Phe Tyr His Ser Pro Ala Asp Asn Lys Arg
                        165   170   175
    Tyr Ile Phe Ala Asp Ala Tyr Ala Gln Tyr Leu Trp Ile Thr Phe Asp
                   180   185   190
    Phe Cys Asn Thr Leu Gln Gly Phe Ser Ile Pro Phe Arg Ala Ala Asp
                   195   200   205
    Leu Leu Leu His Ser Lys Ala Ser Asn Leu Leu Leu Gly Phe Asp Arg
         210   215   220
    Ser His Pro Asn Lys Gln Leu Trp Lys Ser Asp Asp Phe Gly Gln Thr
    225   230   235   240
    Trp Ile Met Ile Gln Glu His Val Lys Ser Phe Ser Trp Gly Ile Asp
</pre>

42

	245			250			255	

Pro Tyr Asp Lys Pro Asn Thr Ile Tyr Ile Glu Arg His Glu Pro Ser

	260			265			270	

Gly Tyr Ser Thr Val Phe Arg Ser Thr Asp Phe Phe Gln Ser Arg Glu

	275			280			285	

Asn Gln Glu Val Ile Leu Glu Glu Val Arg Asp Phe Gln Leu Arg Asp

290   295   300

Lys Tyr Met Phe Ala Thr Lys Val Val His Leu Leu Gly Ser Glu Gln

305   310   315   320

Gln Ser Ser Val Gln Leu Trp Val Ser Phe Gly Arg Lys Pro Met Arg

325   330   335

Ala Ala Gln Phe Val Thr Arg His Pro Ile Asn Glu Tyr Tyr Ile Ala

340   345   350

Asp Ala Ser Glu Asp Gln Val Phe Val Cys Val Ser His Ser Asn Asn

355   360   365

Arg Thr Asn Leu Tyr Ile Ser Glu Ala Glu Gly Leu Lys Phe Ser Leu

370   375   380

Ser Leu Glu Asn Val Leu Tyr Tyr Ser Pro Gly Gly Ala Gly Ser Asp

385   390   395   400

Thr Leu Val Arg Tyr Phe Ala Asn Glu Pro Phe Ala Asp Phe His Arg

405   410   415

Val Glu Gly Leu Gln Gly Val Tyr Ile Ala Thr Leu Ile Asn Gly Ser

420   425   430

Met Asn Glu Glu Asn Met Arg Ser Val Ile Thr Phe Asp Lys Gly Gly

435   440   445

Thr Trp Glu Phe Leu Gln Ala Pro Ala Phe Thr Gly Tyr Gly Glu Lys

450   455   460

Ile Asn Cys Glu Leu Ser Gln Gly Cys Ser Leu His Leu Ala Gln Arg

```
        465                    470                    475                    480
        Leu Ser Gln Leu Leu Asn Leu Gln Leu Arg Arg Met Pro Ile Leu Ser
                        485                    490                    495
        Lys Glu Ser Ala Pro Gly Leu Ile Ile Ala Thr Gly Ser Val Gly Lys
                        500                    505                    510
        Asn Leu Ala Ser Lys Thr Asn Val Tyr Ile Ser Ser Ser Ala Gly Ala
                        515                    520                    525
        Arg Trp Arg Glu Ala Leu Pro Gly Pro His Tyr Tyr Thr Trp Gly Asp
                        530                    535                    540
        His Gly Gly Ile Ile Thr Ala Ile Ala Gln Gly Met Glu Thr Asn Glu
        545                    550                    555                    560
        Leu Lys Tyr Ser Thr Asn Glu Gly Glu Thr Trp Lys Thr Phe Ile Phe
                        565                    570                    575
        Ser Glu Lys Pro Val Phe Val Tyr Gly Leu Leu Thr Glu Pro Gly Glu
                        580                    585                    590
        Lys Ser Thr Val Phe Thr Ile Phe Gly Ser Asn Lys Glu Asn Val His
                        595                    600                    605
        Ser Trp Leu Ile Leu Gln Val Asn Ala Thr Asp Ala Leu Gly Val Pro
                        610                    615                    620
        Cys Thr Glu Asn Asp Tyr Lys Leu Trp Ser Pro Ser Asp Glu Arg Gly
        625                    630                    635                    640
        Asn Glu Cys Leu Leu Gly His Lys Thr Val Phe Lys Arg Arg Thr Pro
                        645                    650                    655
        His Ala Thr Cys Phe Asn Gly Glu Asp Phe Asp Arg Pro Val Val Val
                        660                    665                    670
        Ser Asn Cys Ser Cys Thr Arg Glu Asp Tyr Glu Cys Asp Phe Gly Phe
                        675                    680                    685
        Lys Met Ser Glu Asp Leu Ser Leu Glu Val Cys Val Pro Asp Pro Glu
```

44

690   695   700

Phe Ser Gly Lys Ser Tyr Ser Pro Pro Val Pro Cys Pro Val Gly Ser

705   710   715   720

Thr Tyr Arg Arg Thr Arg Gly Tyr Arg Lys Ile Ser Gly Asp Thr Cys

725   730   735

Ser Gly Gly Asp Val Glu Ala Arg Leu Glu Gly Glu Leu Val Pro Cys

740   745   750

Pro Leu Ala Glu Glu Asn Glu Phe Ile Leu Tyr Ala Val Arg Lys Ser

755   760   765

Ile Tyr Arg Tyr Asp Leu Ala Ser Gly Ala Thr Glu Gln Leu Pro Leu

770   775   780

Thr Gly Leu Arg Ala Ala Val Ala Leu Asp Phe Asp Tyr Glu His Asn

785   790   795   800

Cys Leu Tyr Trp Ser Asp Leu Ala Leu Asp Val Ile Gln Arg Leu Cys

805   810   815

Leu Asn Gly Ser Thr Gly Gln Glu Val Ile Ile Asn Ser Gly Leu Glu

820   825   830

Thr Val Glu Ala Leu Ala Phe Glu Pro Leu Ser Gln Leu Leu Tyr Trp

835   840   845

Val Asp Ala Gly Phe Lys Lys Ile Glu Val Ala Asn Pro Asp Gly Asp

850   855   860

Phe Arg Leu Thr Ile Val Asn Ser Ser Val Leu Asp Arg Pro Arg Ala

865   870   875   880

Leu Val Leu Val Pro Gln Glu Gly Val Met Phe Trp Thr Asp Trp Gly

885   890   895

Asp Leu Lys Pro Gly Ile Tyr Arg Ser Asn Met Asp Gly Ser Ala Ala

900   905   910

Tyr His Leu Val Ser Glu Asp Val Lys Trp Pro Asn Gly Ile Ser Val

915   920   925

Asp Asp Gln Trp Ile Tyr Trp Thr Asp Ala Tyr Leu Glu Cys Ile Glu

930   935   940

Arg Ile Thr Phe Ser Gly Gln Gln Arg Ser Val Ile Leu Asp Asn Leu

945   950   955   960

Pro His Pro Tyr Ala Ile Ala Val Phe Lys Asn Glu Ile Tyr Trp Asp

965   970   975

Asp Trp Ser Gln Leu Ser Ile Phe Arg Ala Ser Lys Tyr Ser Gly Ser

980   985   990

Gln Met Glu Ile Leu Ala Asn Gln Leu Thr Gly Leu Met Asp Met Lys

995   1000   1005

Ile Phe Tyr Lys Gly Lys Asn Thr Gly Ser Asn Ala Cys Val Pro Arg

1010   1015   1020

Pro Cys Ser Leu Leu Cys Leu Pro Lys Ala Asn Asn Ser Arg Ser Cys

1025   1030   1035   1040

Arg Cys Pro Glu Asp Val Ser Ser Ser Val Leu Pro Ser Gly Asp Leu

1045   1050   1055

Met Cys Asp Cys Pro Gln Gly Tyr Gln Leu Lys Asn Asn Thr Cys Val

1060   1065   1070

Lys Glu Glu Asn Thr Cys Leu Arg Asn Gln Tyr Arg Cys Ser Asn Gly

1075   1080   1085

Asn Cys Ile Asn Ser Ile Trp Trp Cys Asp Phe Asp Asn Asp Cys Gly

1090   1095   1100

Asp Met Ser Asp Glu Arg Asn Cys Pro Thr Thr Ile Cys Asp Leu Asp

1105   1110   1115   1120

Thr Gln Phe Arg Cys Gln Glu Ser Gly Thr Cys Ile Pro Leu Ser Tyr

1125   1130   1135

Lys Cys Asp Leu Glu Asp Asp Cys Gly Asp Asn Ser Asp Glu Ser His

1140   1145   1150
Cys Glu Met His Gln Cys Arg Ser Asp Glu Tyr Asn Cys Ser Ser Gly

        1155   1160   1165
Met Cys Ile Arg Ser Ser Trp Val Cys Asp Gly Asp Asn Asp Cys Arg

    1170   1175   1180
Asp Trp Ser Asp Glu Ala Asn Cys Thr Ala Ile Tyr His Thr Cys Glu

1185   1190   1195   1200
Ala Ser Asn Phe Gln Cys Arg Asn Gly His Cys Ile Pro Gln Arg Trp

            1205   1210   1215
Ala Cys Asp Gly Asp Thr Asp Cys Gln Asp Gly Ser Asp Glu Asp Pro

        1220   1225   1230
Val Asn Cys Glu Lys Lys Cys Asn Gly Phe Arg Cys Pro Asn Gly Thr

        1235   1240   1245
Cys Ile Pro Ser Ser Lys His Cys Asp Gly Leu Arg Asp Cys Ser Asp

    1250   1255   1260
Gly Ser Asp Glu Gln His Cys Glu Pro Leu Cys Thr His Phe Met Asp

1265   1270   1275   1280
Phe Val Cys Lys Asn Arg Gln Gln Cys Leu Phe His Ser Met Val Cys

            1285   1290   1295
Asp Gly Ile Ile Gln Cys Arg Asp Gly Ser Asp Glu Asp Ala Ala Phe

        1300   1305   1310
Ala Gly Cys Ser Gln Asp Pro Glu Phe His Lys Val Cys Asp Glu Phe

    1315   1320   1325
Gly Phe Gln Cys Gln Asn Gly Val Cys Ile Ser Leu Ile Trp Lys Cys

        1330   1335   1340
Asp Gly Met Asp Asp Cys Gly Asp Tyr Ser Asp Glu Ala Asn Cys Glu

1345   1350   1355   1360
Asn Pro Thr Glu Ala Pro Asn Cys Ser Arg Tyr Phe Gln Phe Arg Cys

```
                1365                      1370                        1375
      Glu Asn Gly His Cys Ile Pro Asn Arg Trp Lys Cys Asp Arg Glu Asn
                    1380                      1385                    1390
      Asp Cys Gly Asp Trp Ser Asp Glu Lys Asp Cys Gly Asp Ser His Ile
                    1395                      1400                    1405
      Leu Pro Phe Ser Thr Pro Gly Pro Ser Thr Cys Leu Pro Asn Tyr Tyr
              1410                      1415                      1420
      Arg Cys Ser Ser Gly Thr Cys Val Met Asp Thr Trp Val Cys Asp Gly
          1425                      1430                      1435                      1440
      Tyr Arg Asp Cys Ala Asp Gly Ser Asp Glu Glu Ala Cys Pro Leu Leu
                      1445                      1450                    1455
      Ala Asn Val Thr Ala Ala Ser Thr Pro Thr Gln Leu Gly Arg Cys Asp
                  1460                      1465                    1470
      Arg Phe Glu Phe Glu Cys His Gln Pro Lys Thr Cys Ile Pro Asn Trp
                1475                      1480                      1485
      Lys Arg Cys Asp Gly His Gln Asp Cys Gln Asp Gly Arg Asp Glu Ala
              1490                      1495                      1500
      Asn Cys Pro Thr His Ser Thr Leu Thr Cys Met Ser Arg Glu Phe Gln
          1505                      1510                      1515                      1520
      Cys Glu Asp Gly Glu Ala Cys Ile Val Leu Ser Glu Arg Cys Asp Gly
                      1525                      1530                      1535
      Phe Leu Asp Cys Ser Asp Glu Ser Asp Glu Lys Ala Cys Ser Asp Glu
                  1540                      1545                      1550
      Leu Thr Val Tyr Lys Val Gln Asn Leu Gln Trp Thr Ala Asp Phe Ser
                  1555                      1560                      1565
      Gly Asp Val Thr Leu Thr Trp Met Arg Pro Lys Lys Met Pro Ser Ala
              1570                      1575                      1580
      Ser Cys Val Tyr Asn Val Tyr Tyr Arg Val Val Gly Glu Ser Ile Trp
```

```
        1585                 1590                1595                1600
        Lys Thr Leu Glu Thr His Ser Asn Lys Thr Asn Thr Val Leu Lys Val
                        1605                1610                1615
        Leu Lys Pro Asp Thr Thr Tyr Gln Val Lys Val Gln Val Gln Cys Leu
                        1620                1625                1630
        Ser Lys Ala His Asn Thr Asn Asp Phe Val Thr Leu Arg Thr Pro Glu
                    1635                1640                1645
        Gly Leu Pro Asp Ala Pro Arg Asn Leu Gln Leu Ser Leu Pro Arg Glu
                    1650                1655                1660
        Ala Glu Gly Val Ile Val Gly His Trp Ala Pro Pro Ile His Thr His
        1665                1670                1675                1680
        Gly Leu Ile Arg Glu Tyr Ile Val Glu Tyr Ser Arg Ser Gly Ser Lys
                        1685                1690                1695
        Met Trp Ala Ser Gln Arg Ala Ala Ser Asn Phe Thr Glu Ile Lys Asn
                    1700                1705                1710
        Leu Leu Val Asn Thr Leu Tyr Thr Val Arg Val Ala Ala Val Thr Ser
                    1715                1720                1725
        Arg Gly Ile Gly Asn Trp Ser Asp Ser Lys Ser Ile Thr Thr Ile Lys
                    1730                1735                1740
        Gly Lys Val Ile Pro Pro Pro Asp Ile His Ile Asp Ser Tyr Gly Glu
        1745                1750                1755                1760
        Asn Tyr Leu Ser Phe Thr Leu Thr Met Glu Ser Asp Ile Lys Val Asn
                        1765                1770                1775
        Gly Tyr Val Val Asn Leu Phe Trp Ala Phe Asp Thr His Lys Gln Glu
                    1780                1785                1790
        Arg Arg Thr Leu Asn Phe Arg Gly Ser Ile Leu Ser His Lys Val Gly
                    1795                1800                1805
        Asn Leu Thr Ala His Thr Ser Tyr Glu Ile Ser Ala Trp Ala Lys Thr
```

```
        1810                    1815                    1820
   Asp Leu Gly Asp Ser Pro Leu Ala Phe Glu His Val Met Thr Arg Gly
        1825                    1830           1835                   1840
   Val Arg Pro Pro Ala Pro Ser Leu Lys Ala Lys Ala Ile Asn Gln Thr
                        1845                  1850                  1855
   Ala Val Glu Cys Thr Trp Thr Gly Pro Arg Asn Val Val Tyr Gly Ile
                   1860                  1865                  1870
   Phe Tyr Ala Thr Ser Phe Leu Asp Leu Tyr Arg Asn Pro Lys Ser Leu
                   1875                  1880                  1885
   Thr Thr Ser Leu His Asn Lys Thr Val Ile Val Ser Lys Asp Glu Gln
              1890                  1895                  1900
   Tyr Leu Phe Leu Val Arg Val Val Val Pro Tyr Gln Gly Pro Ser Ser
        1905                  1910                  1915                   1920
   Asp Tyr Val Val Val Lys Met Ile Pro Asp Ser Arg Leu Pro Pro Arg
                   1925                  1930                    1935
   His Leu His Val Val His Thr Gly Lys Thr Ser Val Val Ile Lys Trp
                   1940                  1945                  1950
   Glu Ser Pro Tyr Asp Ser Pro Asp Gln Asp Leu Leu Tyr Ala Ile Ala
              1955                  1960                  1965
   Val Lys Asp Leu Ile Arg Lys Thr Asp Arg Ser Tyr Lys Val Lys Ser
        1970                  1975                  1980
   Arg Asn Ser Thr Val Glu Tyr Thr Leu Asn Lys Leu Glu Pro Gly Gly
        1985                  1990                  1995                  2000
   Lys Tyr His Ile Ile Val Gln Leu Gly Asn Met Ser Lys Asp Ser Ser
                   2005                  2010                  2015
   Ile Lys Ile Thr Thr Val Ser Leu Ser Ala Pro Asp Ala Leu Lys Ile
                   2020                  2025                  2030
   Ile Thr Glu Asn Asp His Val Leu Leu Phe Trp Lys Ser Leu Ala Leu
```

	2035					2040					2045			
Lys	Glu	Lys	His	Phe	Asn	Glu	Ser	Arg	Gly	Tyr	Glu	Ile	His	Met	Phe

	2050				2055				2060		
Asp	Ser	Ala	Met	Asn	Ile	Thr	Ala	Tyr	Leu	Gly	Asn	Thr	Thr	Asp	Asn

2065   2070   2075   2080

Phe Phe Lys Ile Ser Asn Leu Lys Met Gly His Asn Tyr Thr Phe Thr

                2085   2090   2095

Val Gln Ala Arg Cys Leu Phe Gly Asn Gln Ile Cys Gly Glu Pro Ala

                2100   2105   2110

Ile Leu Leu Tyr Asp Glu Leu Gly Ser Gly Ala Asp Ala Ser Ala Thr

                2115   2120   2125

Gln Ala Ala Arg Ser Thr Asp Val Ala Ala Val Val Val Pro Ile Leu

                2130   2135   2140

Phe Leu Ile Leu Leu Ser Leu Gly Val Gly Phe Ala Ile Leu Tyr Thr

2145   2150   2155   2160

Lys His Arg Arg Leu Gln Ser Ser Phe Thr Ala Phe Ala Asn Ser His

                2165   2170   2175

Tyr Ser Ser Arg Leu Gly Ser Ala Ile Phe Ser Ser Gly Asp Asp Leu

                2180   2185   2190

Gly Glu Asp Asp Glu Asp Ala Pro Met Ile Thr Gly Phe Ser Asp Asp

                2195   2200   2205

Val Pro Met Val Ile Ala

2210

Sequence ID No. 7

Length of the Sequence: 6843

Type: nucleic acid

Strandedness: double

Topology: linear

51

Molecular type: cDNA to mRNA

Feature:

 Name/Key: sig peptide

 Location: 81..164

 Identification method: S

 Name/Key: mat peptide

 Location: 165..6722

 Identification method: S

Sequence:

```
                                        CCG GCCCAGCGGC TCTCCTGGCC     23

TCGCGCTGCA CATTCTCTCC TGGCGGCGGC GCCACCTGCA GTAGCGTTCG CCCGAACATG     83
                                                          Met
                                                            1
GCG ACA CGG AGC AGC AGG AGG GAG TCG CGA CTC CCG TTC CTA TTC ACC      131
Ala Thr Arg Ser Ser Arg Arg Glu Ser Arg Leu Pro Phe Leu Phe Thr
             5                  10                  15
CTG GTC GCA CTG CTG CCG CCC GGA GCT CTC TGC GAA GTC TGG ACG CAG      179
Leu Val Ala Leu Leu Pro Pro Gly Ala Leu Cys Glu Val Trp Thr Gln
             20                  25                  30
AGG CTG CAC GGC GGC AGC GCG CCC TTG CCC CAG GAC CGG GGC TTC CTC      227
Arg Leu His Gly Gly Ser Ala Pro Leu Pro Gln Asp Arg Gly Phe Leu
             35                  40                  45
GTG GTG CAG GGC GAC CCG CGC GAG CTG CGG CTG TGG GCG CGC GGG GAT      275
Val Val Gln Gly Asp Pro Arg Glu Leu Arg Leu Trp Ala Arg Gly Asp
    50                  55                  60                  65
GCC AGG GGG GCG AGC CGC GCG GAC GAG AAG CCG CTC CGG AGG AAA CGG      323
Ala Arg Gly Ala Ser Arg Ala Asp Glu Lys Pro Leu Arg Arg Lys Arg
             70                  75                  80
```

AGC GCT GCC CTG CAG CCC GAG CCC ATC AAG GTG TAC GGA CAG GTT AGT   371
Ser Ala Ala Leu Gln Pro Glu Pro Ile Lys Val Tyr Gly Gln Val Ser
              85   90   95

CTG AAT GAT TCC CAC AAT CAG ATG GTG GTG CAC TGG GCT GGA GAG AAA   419
Leu Asn Asp Ser His Asn Gln Met Val Val His Trp Ala Gly Glu Lys
             100   105   110

AGC AAC GTG ATC GTG GCC TTG GCC CGA GAT AGC CTG GCA TTG GCG AGG   467
Ser Asn Val Ile Val Ala Leu Ala Arg Asp Ser Leu Ala Leu Ala Arg
             115   120   125

CCC AAG AGC AGT GAT GTG TAC GTG TCT TAC GAC TAT GGA AAA TCA TTC   515
Pro Lys Ser Ser Asp Val Tyr Val Ser Tyr Asp Tyr Gly Lys Ser Phe
130   135   140   145

AAG AAA ATT TCA GAC AAG TTA AAC TTT GGC TTG GGA AAT AGG AGT GAA   563
Lys Lys Ile Ser Asp Lys Leu Asn Phe Gly Leu Gly Asn Arg Ser Glu
             150   155   160

GCT GTT ATC GCC CAG TTC TAC CAC AGC CCT GCG GAC AAC AAG CGG TAC   611
Ala Val Ile Ala Gln Phe Tyr His Ser Pro Ala Asp Asn Lys Arg Tyr
             165   170   175

ATC TTT GCA GAC GCT TAT GCC CAG TAC CTC TGG ATC ACG TTT GAC TTC   659
Ile Phe Ala Asp Ala Tyr Ala Gln Tyr Leu Trp Ile Thr Phe Asp Phe
             180   185   190

TGC AAC ACT CTT CAA GGC TTT TCC ATC CCA TTT CGG GCA GCT GAT CTC   707
Cys Asn Thr Leu Gln Gly Phe Ser Ile Pro Phe Arg Ala Ala Asp Leu
             195   200   205

CTC CTA CAC AGT AAG GCC TCC AAC CTT CTC TTG GGC TTT GAC AGG TCC   755
Leu Leu His Ser Lys Ala Ser Asn Leu Leu Leu Gly Phe Asp Arg Ser
210   215   220   225

CAC CCC AAC AAG CAG CTG TGG AAG TCA GAT GAC TTT GGC CAG ACC TGG   803

His Pro Asn Lys Gln Leu Trp Lys Ser Asp Asp Phe Gly Gln Thr Trp
            230   235   240

ATC ATG ATT CAG GAA CAT GTC AAG TCC TTT TCT TGG GGA ATT GAT CCC   851
Ile Met Ile Gln Glu His Val Lys Ser Phe Ser Trp Gly Ile Asp Pro
            245   250   255

TAT GAC AAA CCA AAT ACC ATC TAC ATT GAA CGA CAC GAA CCC TCT GGC   899
Tyr Asp Lys Pro Asn Thr Ile Tyr Ile Glu Arg His Glu Pro Ser Gly
          260   265   270

TAC TCC ACT GTC TTC CGA AGT ACA GAT TTC TTC CAG TCC CGG GAA AAC   947
Tyr Ser Thr Val Phe Arg Ser Thr Asp Phe Phe Gln Ser Arg Glu Asn
      275   280   285

CAG GAA GTG ATC CTT GAG GAA GTG AGA GAT TTT CAG CTT CGG GAC AAG   995
Gln Glu Val Ile Leu Glu Glu Val Arg Asp Phe Gln Leu Arg Asp Lys
290   295   300   305

TAC ATG TTT GCT ACA AAG GTG GTG CAT CTC TTG GGC AGT GAA CAG CAG   1043
Tyr Met Phe Ala Thr Lys Val Val His Leu Leu Gly Ser Glu Gln Gln
            310   315   320

TCT TCT GTC CAG CTC TGG GTC TCC TTT GGC CGG AAG CCC ATG AGA GCA   1091
Ser Ser Val Gln Leu Trp Val Ser Phe Gly Arg Lys Pro Met Arg Ala
            325   330   335

GCC CAG TTT GTC ACA AGA CAT CCT ATT AAT GAA TAT TAC ATC GCA GAT   1139
Ala Gln Phe Val Thr Arg His Pro Ile Asn Glu Tyr Tyr Ile Ala Asp
            340   345   350

GCC TCC GAG GAC CAG GTG TTT GTG TGT GTC AGC CAC AGT AAC AAC CGC   1187
Ala Ser Glu Asp Gln Val Phe Val Cys Val Ser His Ser Asn Asn Arg
          355   360   365

ACC AAT TTA TAC ATC TCA GAG GCA GAG GGG CTG AAG TTC TCC CTG TCC   1235
Thr Asn Leu Tyr Ile Ser Glu Ala Glu Gly Leu Lys Phe Ser Leu Ser

```
           370                      375                    380                     385
          TTG GAG AAC GTG CTC TAT TAC AGC CCA GGA GGG GCC GGC AGT GAC ACC      1283
          Leu Glu Asn Val Leu Tyr Tyr Ser Pro Gly Gly Ala Gly Ser Asp Thr

                             390                    395                    400
          TTG GTG AGG TAT TTT GCA AAT GAA CCA TTT GCT GAC TTC CAC CGA GTG      1331
          Leu Val Arg Tyr Phe Ala Asn Glu Pro Phe Ala Asp Phe His Arg Val

                        405                    410                    415
          GAA GGA TTG CAA GGA GTC TAC ATT GCT ACT CTG ATT AAT GGT TCT ATG      1379
          Glu Gly Leu Gln Gly Val Tyr Ile Ala Thr Leu Ile Asn Gly Ser Met

                        420                    425                    430
          AAT GAG GAG AAC ATG AGA TCG GTC ATC ACC TTT GAC AAA GGG GGA ACC      1427
          Asn Glu Glu Asn Met Arg Ser Val Ile Thr Phe Asp Lys Gly Gly Thr

                   435                    440                    445
          TGG GAG TTT CTT CAG GCT CCA GCC TTC ACG GGA TAT GGA GAG AAA ATC      1475
          Trp Glu Phe Leu Gln Ala Pro Ala Phe Thr Gly Tyr Gly Glu Lys Ile

          450                    455                    460                    465
          AAT TGT GAG CTT TCC CAG GGC TGT TCC CTT CAT CTG GCT CAG CGC CTC      1523
          Asn Cys Glu Leu Ser Gln Gly Cys Ser Leu His Leu Ala Gln Arg Leu

                             470                    475                    480
          AGT CAG CTC CTC AAC CTC CAG CTC CGG AGA ATG CCC ATC CTG TCC AAG      1571
          Ser Gln Leu Leu Asn Leu Gln Leu Arg Arg Met Pro Ile Leu Ser Lys

                        485                    490                    495
          GAG TCG GCT CCA GGC CTC ATC ATC GCC ACT GGC TCA GTG GGA AAG AAC      1619
          Glu Ser Ala Pro Gly Leu Ile Ile Ala Thr Gly Ser Val Gly Lys Asn

                        500                    505                    510
          TTG GCT AGC AAG ACA AAC GTG TAC ATC TCT AGC AGT GCT GGA GCC AGG      1667
          Leu Ala Ser Lys Thr Asn Val Tyr Ile Ser Ser Ser Ala Gly Ala Arg

                   515                    520                    525
```

TGG CGA GAG GCA CTT CCT GGA CCT CAC TAC TAC ACA TGG GGA GAC CAC   1715
Trp Arg Glu Ala Leu Pro Gly Pro His Tyr Tyr Thr Trp Gly Asp His
530   535   540   545

GGC GGA ATC ATC ACG GCC ATT GCC CAG GGC ATG GAA ACC AAC GAG CTA   ·1763
Gly Gly Ile Ile Thr Ala Ile Ala Gln Gly Met Glu Thr Asn Glu Leu
                550   555   560

AAA TAC AGT ACC AAT GAA GGG GAG ACC TGG AAA ACA TTC ATC TTC TCT   1811
Lys Tyr Ser Thr Asn Glu Gly Glu Thr Trp Lys Thr Phe Ile Phe Ser
                565   570   575

GAG AAG CCA GTG TTT GTG TAT GGC CTC CTC ACA GAA CCT GGG GAG AAG   1859
Glu Lys Pro Val Phe Val Tyr Gly Leu Leu Thr Glu Pro Gly Glu Lys
                580   585   590

AGC ACT GTC TTC ACC ATC TTT GGC TCG AAC AAA GAG AAT GTC CAC AGC   1907
Ser Thr Val Phe Thr Ile Phe Gly Ser Asn Lys Glu Asn Val His Ser
       595   600   605

TGG CTG ATC CTC CAG GTC AAT GCC ACG GAT GCC TTG GGA GTT CCC TGC   1955
Trp Leu Ile Leu Gln Val Asn Ala Thr Asp Ala Leu Gly Val Pro Cys
610   615   620   625

ACA GAG AAT GAC TAC AAG CTG TGG TCA CCA TCT GAT GAG CGG GGG AAT   2003
Thr Glu Asn Asp Tyr Lys Leu Trp Ser Pro Ser Asp Glu Arg Gly Asn
                630   635   640

GAG TGT TTG CTG GGA CAC AAG ACT GTT TTC AAA CGG CGG ACC CCC CAT   2051
Glu Cys Leu Leu Gly His Lys Thr Val Phe Lys Arg Arg Thr Pro His
                645   650   655

GCC ACA TGC TTC AAT GGA GAG GAC TTT GAC AGG CCG GTG GTC GTG TCC   2099
Ala Thr Cys Phe Asn Gly Glu Asp Phe Asp Arg Pro Val Val Val Ser
                660   665   670

AAC TGC TCC TGC ACC CGG GAG GAC TAT GAG TGT GAC TTC GGT TTC AAG   2147

55

56

Asn Cys Ser Cys Thr Arg Glu Asp Tyr Glu Cys Asp Phe Gly Phe Lys

   675   680   685

ATG AGT GAA GAT TTG TCA TTA GAG GTT TGT GTT CCA GAT CCG GAA TTT   2195

Met Ser Glu Asp Leu Ser Leu Glu Val Cys Val Pro Asp Pro Glu Phe

690   695   700   705

TCT GGA AAG TCA TAC TCC CCT CCT GTG CCT TGC CCT GTG GGT TCT ACT   2243

Ser Gly Lys Ser Tyr Ser Pro Pro Val Pro Cys Pro Val Gly Ser Thr

       710   715   720

TAC AGG AGA ACG AGA GGC TAC CGG AAG ATT TCT GGG GAC ACT TGT AGC   2291

Tyr Arg Arg Thr Arg Gly Tyr Arg Lys Ile Ser Gly Asp Thr Cys Ser

       725   730   735

GGA GGA GAT GTT GAA GCG CGA CTG GAA GGA GAG CTG GTC CCC TGT CCC   2339

Gly Gly Asp Val Glu Ala Arg Leu Glu Gly Glu Leu Val Pro Cys Pro

       740   745   750

CTG GCA GAA GAG AAC GAG TTC ATT CTG TAT GCT GTG AGG AAA TCC ATC   2387

Leu Ala Glu Glu Asn Glu Phe Ile Leu Tyr Ala Val Arg Lys Ser Ile

   755   760   765

TAC CGC TAT GAC CTG GCC TCG GGA GCC ACC GAG CAG TTG CCT CTC ACC   2435

Tyr Arg Tyr Asp Leu Ala Ser Gly Ala Thr Glu Gln Leu Pro Leu Thr

770   775   780   785

GGG CTA CGG GCA GCA GTG GCC CTG GAC TTT GAC TAT GAG CAC AAC TGT   2483

Gly Leu Arg Ala Ala Val Ala Leu Asp Phe Asp Tyr Glu His Asn Cys

       790   795   800

TTG TAT TGG TCC GAC CTG GCC TTG GAC GTC ATC CAG CGC CTC TGT TTG   2531

Leu Tyr Trp Ser Asp Leu Ala Leu Asp Val Ile Gln Arg Leu Cys Leu

       805   810   815

AAT GGA AGC ACA GGG CAA GAG GTG ATC ATC AAT TCT GGC CTG GAG ACA   2579

Asn Gly Ser Thr Gly Gln Glu Val Ile Ile Asn Ser Gly Leu Glu Thr

```
                820                   825                   830
        GTA GAA GCT TTG GCT TTT GAA CCC CTC AGC CAG CTG CTT TAC TGG GTA      2627
        Val Glu Ala Leu Ala Phe Glu Pro Leu Ser Gln Leu Leu Tyr Trp Val
                835                   840                   845
        GAT GCA GGC TTC AAA AAG ATT GAG GTA GCT AAT CCA GAT GGC GAC TTC      2675
        Asp Ala Gly Phe Lys Lys Ile Glu Val Ala Asn Pro Asp Gly Asp Phe
        850                   855                   860                   865
        CGA CTC ACA ATC GTC AAT TCC TCT GTG CTT GAT CGT CCC AGG GCT CTG      2723
        Arg Leu Thr Ile Val Asn Ser Ser Val Leu Asp Arg Pro Arg Ala Leu
                        870                   875                   880
        GTC CTC GTG CCC CAA GAG GGG GTG ATG TTC TGG ACA GAC TGG GGA GAC      2771
        Val Leu Val Pro Gln Glu Gly Val Met Phe Trp Thr Asp Trp Gly Asp
                        885                   890                   895
        CTG AAG CCT GGG ATT TAT CGG AGC AAT ATG GAT GGT TCT GCT GCC TAT      2819
        Leu Lys Pro Gly Ile Tyr Arg Ser Asn Met Asp Gly Ser Ala Ala Tyr
                    900                   905                   910
        CAC CTG GTG TCT GAG GAT GTG AAG TGG CCC AAT GGC ATC TCT GTG GAC      2867
        His Leu Val Ser Glu Asp Val Lys Trp Pro Asn Gly Ile Ser Val Asp
            915                   920                   925
        GAC CAG TGG ATT TAC TGG ACG GAT GCC TAC CTG GAG TGC ATA GAG CGG      2915
        Asp Gln Trp Ile Tyr Trp Thr Asp Ala Tyr Leu Glu Cys Ile Glu Arg
        930                   935                   940                   945
        ATC ACG TTC AGT GGC CAG CAG CGC TCT GTC ATT CTG GAC AAC CTC CCG      2963
        Ile Thr Phe Ser Gly Gln Gln Arg Ser Val Ile Leu Asp Asn Leu Pro
                    950                   955                   960
        CAC CCC TAT GCC ATT GCT GTC TTT AAG AAT GAA ATC TAC TGG GAT GAC      3011
        His Pro Tyr Ala Ile Ala Val Phe Lys Asn Glu Ile Tyr Trp Asp Asp
                965                   970                   975
```

TGG TCA CAG CTC AGC ATA TTC CGA GCT TCC AAA TAC AGT GGG TCC CAG   3059
Trp Ser Gln Leu Ser Ile Phe Arg Ala Ser Lys Tyr Ser Gly Ser Gln
        980   985   990

ATG GAG ATT CTG GCA AAC CAG CTC ACG GGG CTC ATG GAC ATG AAG ATT   ·3107
Met Glu Ile Leu Ala Asn Gln Leu Thr Gly Leu Met Asp Met Lys Ile
        995   1000   1005

TTC TAC AAG GGG AAG AAC ACT GGA AGC AAT GCC TGT GTG CCC AGG CCA   ·3155
Phe Tyr Lys Gly Lys Asn Thr Gly Ser Asn Ala Cys Val Pro Arg Pro
1010   1015   1020   1025

TGC AGC CTG CTG TGC CTG CCC AAG GCC AAC AAC AGT AGA AGC TGC AGG   3203
Cys Ser Leu Leu Cys Leu Pro Lys Ala Asn Asn Ser Arg Ser Cys Arg
           1030   1035   1040

TGT CCA GAG GAT GTG TCC AGC AGT GTG CTT CCA TCA GGG GAC CTG ATG   3251
Cys Pro Glu Asp Val Ser Ser Ser Val Leu Pro Ser Gly Asp Leu Met
          1045   1050   1055

TGT GAC TGC CCT CAG GGC TAT CAG CTC AAG AAC AAT ACC TGT GTC AAA   3299
Cys Asp Cys Pro Gln Gly Tyr Gln Leu Lys Asn Asn Thr Cys Val Lys
       1060   1065   1070

GAA GAG AAC ACC TGT CTT CGC AAC CAG TAT CGC TGC AGC AAC GGG AAC   3347
Glu Glu Asn Thr Cys Leu Arg Asn Gln Tyr Arg Cys Ser Asn Gly Asn
     1075   1080   1085

TGT ATC AAC AGC ATT TGG TGG TGT GAC TTT GAC AAC GAC TGT GGA GAC   3395
Cys Ile Asn Ser Ile Trp Trp Cys Asp Phe Asp Asn Asp Cys Gly Asp
1090   1095   1100   1105

ATG AGC GAT GAG AGA AAC TGC CCT ACC ACC ATC TGT GAC CTG GAC ACC   3443
Met Ser Asp Glu Arg Asn Cys Pro Thr Thr Ile Cys Asp Leu Asp Thr
           1110   1115·   1120

CAG TTT CGT TGC CAG GAG TCT GGG ACT TGT ATC CCA CTG TCC TAT AAA   3491

Gln Phe Arg Cys Gln Glu Ser Gly Thr Cys Ile Pro Leu Ser Tyr Lys
             1125   1130   1135

TGT GAC CTT GAG GAT GAC TGT GGA GAC AAC AGT GAT GAA AGT CAT TGT   3539
Cys Asp Leu Glu Asp Asp Cys Gly Asp Asn Ser Asp Glu Ser His Cys
         1140   1145   1150

GAA ATG CAC CAG TGC CGG AGT GAC GAG TAC AAC TGC AGT TCC GGC ATG   3587
Glu Met His Gln Cys Arg Ser Asp Glu Tyr Asn Cys Ser Ser Gly Met
         1155   1160   1165

TGC ATC CGC TCC TCC TGG GTA TGT GAC GGG GAC AAC GAC TGC AGG GAC   3635
Cys Ile Arg Ser Ser Trp Val Cys Asp Gly Asp Asn Asp Cys Arg Asp
1170   1175   1180   1185

TGG TCT GAT GAA GCC AAC TGT ACC GCC ATC TAT CAC ACC TGT GAG GCC   3683
Trp Ser Asp Glu Ala Asn Cys Thr Ala Ile Tyr His Thr Cys Glu Ala
             1190   1195   1200

TCC AAC TTC CAG TGC CGA AAC GGG CAC TGC ATC CCC CAG CGG TGG GCG   3731
Ser Asn Phe Gln Cys Arg Asn Gly His Cys Ile Pro Gln Arg Trp Ala
             1205   1210   1215

TGT GAC GGG GAT ACG GAC TGC CAG GAT GGT TCC GAT GAG GAT CCA GTC   3779
Cys Asp Gly Asp Thr Asp Cys Gln Asp Gly Ser Asp Glu Asp Pro Val
         1220   1225   1230

AAC TGT GAG AAG AAG TGC AAT GGA TTC CGC TGC CCA AAC GGC ACT TGC   3827
Asn Cys Glu Lys Lys Cys Asn Gly Phe Arg Cys Pro Asn Gly Thr Cys
     1235   1240   1245

ATC CCA TCC AGC AAA CAT TGT GAT GGT CTG CGT GAT TGC TCT GAT GGC   3875
Ile Pro Ser Ser Lys His Cys Asp Gly Leu Arg Asp Cys Ser Asp Gly
1250   1255   1260   1265

TCC GAT GAA CAG CAC TGC GAG CCC CTC TGT ACG CAC TTC ATG GAC TTT   3923
Ser Asp Glu Gln His Cys Glu Pro Leu Cys Thr His Phe Met Asp Phe

```
                1270              1275              1280
     GTG TGT AAG AAC CGC CAG CAG TGC CTG TTC CAC TCC ATG GTC TGT GAC     3971
     Val Cys Lys Asn Arg Gln Gln Cys Leu Phe His Ser Met Val Cys Asp
                     1285              1290              1295
     GGA ATC ATC CAG TGC CGC GAC GGG TCC GAT GAG GAT GCG GCG TTT GCA     4019
     Gly Ile Ile Gln Cys Arg Asp Gly Ser Asp Glu Asp Ala Ala Phe Ala
                1300              1305              1310
     GGA TGC TCC CAA GAT CCT GAG TTC CAC AAG GTA TGT GAT GAG TTC GGT     4067
     Gly Cys Ser Gln Asp Pro Glu Phe His Lys Val Cys Asp Glu Phe Gly
                1315              1320              1325
     TTC CAG TGT CAG AAT GGA GTG TGC ATC AGT TTG ATT TGG AAG TGC GAC     4115
     Phe Gln Cys Gln Asn Gly Val Cys Ile Ser Leu Ile Trp Lys Cys Asp
          1330              1335              1340              1345
     GGG ATG GAT GAT TGC GGC GAT TAT TCT GAT GAA GCC AAC TGC GAA AAC     4163
     Gly Met Asp Asp Cys Gly Asp Tyr Ser Asp Glu Ala Asn Cys Glu Asn
                          1350              1355              1360
     CCC ACA GAA GCC CCA AAC TGC TCC CGC TAC TTC CAG TTT CGG TGT GAG     4211
     Pro Thr Glu Ala Pro Asn Cys Ser Arg Tyr Phe Gln Phe Arg Cys Glu
                     1365              1370              1375
     AAT GGC CAC TGC ATC CCC AAC AGA TGG AAA TGT GAC AGG GAG AAC GAC     4259
     Asn Gly His Cys Ile Pro Asn Arg Trp Lys Cys Asp Arg Glu Asn Asp
                1380              1385              1390
     TGT GGG GAC TGG TCT GAT GAG AAG GAT TGT GGA GAT TCA CAT ATT CTT     4307
     Cys Gly Asp Trp Ser Asp Glu Lys Asp Cys Gly Asp Ser His Ile Leu
                1395              1400              1405
     CCC TTC TCG ACT CCT GGG CCC TCC ACG TGT CTG CCC AAT TAC TAC CGC     4355
     Pro Phe Ser Thr Pro Gly Pro Ser Thr Cys Leu Pro Asn Tyr Tyr Arg
                1410              1415              1420              1425
```

TGC AGC AGT GGG ACC TGC GTG ATG GAC ACC TGG GTG TGC GAC GGG TAC   4403
Cys Ser Ser Gly Thr Cys Val Met Asp Thr Trp Val Cys Asp Gly Tyr
       1430   1435   1440

CGA GAT TGT GCA GAT GGC TCT GAC GAG GAA GCC TGC CCC TTG CTT GCA   4451
Arg Asp Cys Ala Asp Gly Ser Asp Glu Glu Ala Cys Pro Leu Leu Ala
       1445   1450   1455

AAC GTC ACT GCT GCC TCC ACT CCC ACC CAA CTT GGG CGA TGT GAC CGA   4499
Asn Val Thr Ala Ala Ser Thr Pro Thr Gln Leu Gly Arg Cys Asp Arg
       1460   1465   1470

TTT GAG TTC GAA TGC CAC CAA CCG AAG ACG TGT ATT CCC AAC TGG AAG   4547
Phe Glu Phe Glu Cys His Gln Pro Lys Thr Cys Ile Pro Asn Trp Lys
    1475   1480   1485

CGC TGT GAC GGC CAC CAA GAT TGC CAG GAT GGC CGG GAC GAG GCC AAT   4595
Arg Cys Asp Gly His Gln Asp Cys Gln Asp Gly Arg Asp Glu Ala Asn
1490   1495   1500   1505

TGC CCC ACA CAC AGC ACC TTG ACT TGC ATG AGC AGG GAG TTC CAG TGC   4643
Cys Pro Thr His Ser Thr Leu Thr Cys Met Ser Arg Glu Phe Gln Cys
       1510   1515   1520

GAG GAC GGG GAG GCC TGC ATT GTG CTC TCG GAG CGC TGC GAC GGC TTC   4691
Glu Asp Gly Glu Ala Cys Ile Val Leu Ser Glu Arg Cys Asp Gly Phe
       1525   1530   1535

CTG GAC TGC TCG GAC GAG AGC GAT GAA AAG GCC TGC AGT GAT GAG TTG   4739
Leu Asp Cys Ser Asp Glu Ser Asp Glu Lys Ala Cys Ser Asp Glu Leu
      1540   1545   1550

ACT GTG TAC AAA GTA CAG AAT CTT CAG TGG ACA GCT GAC TTC TCT GGG   4787
Thr Val Tyr Lys Val Gln Asn Leu Gln Trp Thr Ala Asp Phe Ser Gly
    1555   1560   1565

GAT GTG ACT TTG ACC TGG ATG AGG CCC AAA AAA ATG CCC TCT GCA TCT   4835

62

Asp Val Thr Leu Thr Trp Met Arg Pro Lys Lys Met Pro Ser Ala Ser
1570   1575   1580   1585

TGT GTA TAT AAT GTC TAC TAC AGG GTG GTT GGA GAG AGC ATA TGG AAG   4883
Cys Val Tyr Asn Val Tyr Tyr Arg Val Val Gly Glu Ser Ile Trp Lys
                1590   1595   1600

ACT CTG GAG ACC CAC AGC AAT AAG ACA AAC ACT GTA TTA AAA GTC TTG   4931
Thr Leu Glu Thr His Ser Asn Lys Thr Asn Thr Val Leu Lys Val Leu
                1605   1610   1615

AAA CCA GAT ACC ACG TAT CAG GTT AAA GTA CAG GTT CAG TGT CTC AGC   4979
Lys Pro Asp Thr Thr Tyr Gln Val Lys Val Gln Val Gln Cys Leu Ser
                1620   1625   1630

AAG GCA CAC AAC ACC AAT GAC TTT GTG ACC CTG AGG ACC CCA GAG GGA   5027
Lys Ala His Asn Thr Asn Asp Phe Val Thr Leu Arg Thr Pro Glu Gly
          1635   1640   1645

TTG CCA GAT GCC CCT CGA AAT CTC CAG CTG TCA CTC CCC AGG GAA GCA   5075
Leu Pro Asp Ala Pro Arg Asn Leu Gln Leu Ser Leu Pro Arg Glu Ala
1650   1655   1660   1665

GAA GGT GTG ATT GTA GGC CAC TGG GCT CCT CCC ATC CAC ACC CAT GGC   5123
Glu Gly Val Ile Val Gly His Trp Ala Pro Pro Ile His Thr His Gly
                1670   1675   1680

CTC ATC CGT GAG TAC ATT GTA GAA TAC AGC AGG AGT GGT TCC AAG ATG   5171
Leu Ile Arg Glu Tyr Ile Val Glu Tyr Ser Arg Ser Gly Ser Lys Met
                1685   1690   1695

TGG GCC TCC CAG AGG GCT GCT AGT AAC TTT ACA GAA ATC AAG AAC TTA   5219
Trp Ala Ser Gln Arg Ala Ala Ser Asn Phe Thr Glu Ile Lys Asn Leu
          1700   1705   1710

TTG GTC AAC ACT CTA TAC ACC GTC AGA GTG GCT GCG GTG ACT AGT CGT   5267
Leu Val Asn Thr Leu Tyr Thr Val Arg Val Ala Ala Val Thr Ser Arg

1715   1720   1725

GGA ATA GGA AAC TGG AGC GAT TCT AAA TCC ATT ACC ACC ATA AAA GGA   5315

Gly Ile Gly Asn Trp Ser Asp Ser Lys Ser Ile Thr Thr Ile Lys Gly

1730   1735   1740   1745

AAA GTG ATC CCA CCA CCA GAT ATC CAC ATT GAC AGC TAT GGT GAA AAT   5363

Lys Val Ile Pro Pro Pro Asp Ile His Ile Asp Ser Tyr Gly Glu Asn

            1750   1755   1760

TAT CTA AGC TTC ACC CTG ACC ATG GAG AGT GAT ATC AAG GTG AAT GGC   5411

Tyr Leu Ser Phe Thr Leu Thr Met Glu Ser Asp Ile Lys Val Asn Gly

            1765   1770   1775

TAT GTG GTG AAC CTT TTC TGG GCA TTT GAC ACC CAC AAG CAA GAG AGG   5459

Tyr Val Val Asn Leu Phe Trp Ala Phe Asp Thr His Lys Gln Glu Arg

       1780   1785   1790

AGA ACT TTG AAC TTC CGA GGA AGC ATA TTG TCA CAC AAA GTT GGC AAT   5507

Arg Thr Leu Asn Phe Arg Gly Ser Ile Leu Ser His Lys Val Gly Asn

      1795   1800   1805

CTG ACA GCT CAT ACA TCC TAT GAG ATT TCT GCC TGG GCC AAG ACT GAC   5555

Leu Thr Ala His Thr Ser Tyr Glu Ile Ser Ala Trp Ala Lys Thr Asp

1810   1815   1820   1825

TTG GGG GAT AGC CCT CTG GCA TTT GAG CAT GTT ATG ACC AGA GGG GTT   5603

Leu Gly Asp Ser Pro Leu Ala Phe Glu His Val Met Thr Arg Gly Val

            1830   1835   1840

CGC CCA CCT GCA CCT AGC CTC AAG GCC AAA GCC ATC AAC CAG ACT GCA   5651

Arg Pro Pro Ala Pro Ser Leu Lys Ala Lys Ala Ile Asn Gln Thr Ala

        1845   1850   1855

GTG GAA TGT ACC TGG ACC GGC CCC CGG AAT GTG GTT TAT GGT ATT TTC   5699

Val Glu Cys Thr Trp Thr Gly Pro Arg Asn Val Val Tyr Gly Ile Phe

      1860   1865   1870

TAT GCC ACG TCC TTT CTT GAC CTC TAT CGC AAC CCG AAG AGC TTG ACT · 5747

Tyr Ala Thr Ser Phe Leu Asp Leu Tyr Arg Asn Pro Lys Ser Leu Thr

    1875   1880   1885

ACT TCA CTC CAC AAC AAG ACG GTC ATT GTC AGT AAG GAT GAG CAG TAT   5795

Thr Ser Leu His Asn Lys Thr Val Ile Val Ser Lys Asp Glu Gln Tyr

1890   1895   1900   1905

TTG TTT CTG GTC CGT GTA GTG GTA CCC TAC CAG GGG CCA TCC TCT GAC   5843

Leu Phe Leu Val Arg Val Val Val Pro Tyr Gln Gly Pro Ser Ser Asp

        1910   1915   1920

TAC GTT GTA GTG AAG ATG ATC CCG GAC AGC AGG CTT CCA CCC CGT CAC   5891

Tyr Val Val Val Lys Met Ile Pro Asp Ser Arg Leu Pro Pro Arg His

     1925   1930   1935

CTG CAT GTG GTT CAT ACG GGC AAA ACC TCC GTG GTC ATC AAG TGG GAA   5939

Leu His Val Val His Thr Gly Lys Thr Ser Val Val Ile Lys Trp Glu

    1940   1945   1950

TCA CCG TAT GAC TCT CCT GAC CAG GAC TTG TTG TAT GCA ATT GCA GTC   5987

Ser Pro Tyr Asp Ser Pro Asp Gln Asp Leu Leu Tyr Ala Ile Ala Val

    1955   1960   1965

AAA GAT CTC ATA AGA AAG ACT GAC AGG AGC TAC AAA GTA AAA TCC CGT   6035

Lys Asp Leu Ile Arg Lys Thr Asp Arg Ser Tyr Lys Val Lys Ser Arg

1970   1975   1980   1985

AAC AGC ACT GTG GAA TAC ACC CTT AAC AAG TTG GAG CCT GGC GGG AAA   6083

Asn Ser Thr Val Glu Tyr Thr Leu Asn Lys Leu Glu Pro Gly Gly Lys

       1990   1995   2000

TAC CAC ATC ATT GTC CAA CTG GGG AAC ATG AGC AAA GAT TCC AGC ATA   6131

Tyr His Ile Ile Val Gln Leu Gly Asn Met Ser Lys Asp Ser Ser Ile

       2005   2010   2015

AAA ATT ACC ACA GTT TCA TTA TCA GCA CCT GAT GCC TTA AAA ATC ATA   6179

Lys Ile Thr Thr Val Ser Leu Ser Ala Pro Asp Ala Leu Lys Ile Ile
     2020   2025   2030

ACA GAA AAT GAT CAT GTT CTT CTG TTT TGG AAA AGC CTG GCT TTA AAG   6227
Thr Glu Asn Asp His Val Leu Leu Phe Trp Lys Ser Leu Ala Leu Lys
     2035   2040   2045

GAA AAG CAT TTT AAT GAA AGC AGG GGC TAT GAG ATA CAC ATG TTT GAT   6275
Glu Lys His Phe Asn Glu Ser Arg Gly Tyr Glu Ile His Met Phe Asp
2050   2055   2060   2065

AGT GCC ATG AAT ATC ACA GCT TAC CTT GGG AAT ACT ACT GAC AAT TTC   6323 .
Ser Ala Met Asn Ile Thr Ala Tyr Leu Gly Asn Thr Thr Asp Asn Phe
        2070   2075   2080

TTT AAA ATT TCC AAC CTG AAG ATG GGT CAT AAT TAC ACG TTC ACC GTC   6371
Phe Lys Ile Ser Asn Leu Lys Met Gly His Asn Tyr Thr Phe Thr Val
      2085   2090   2095

CAA GCA AGA TGC CTT TTT GGC AAC CAG ATC TGT GGG GAG CCT GCC ATC   6419
Gln Ala Arg Cys Leu Phe Gly Asn Gln Ile Cys Gly Glu Pro Ala Ile
     2100   2105   2110

CTG CTG TAC GAT GAG CTG GGG TCT GGT GCA GAT GCA TCT GCA ACG CAG   6467
Leu Leu Tyr Asp Glu Leu Gly Ser Gly Ala Asp Ala Ser Ala Thr Gln
    2115   2120   2125

GCT GCC AGA TCT ACG GAT GTT GCT GCT GTG GTG GTG CCC ATC TTA TTC   6515
Ala Ala Arg Ser Thr Asp Val Ala Ala Val Val Val Pro Ile Leu Phe
2130   2135   2140   2145

CTG ATA CTG CTG AGC CTG GGG GTG GGG TTT GCC ATC CTG TAC ACG AAG   6563
Leu Ile Leu Leu Ser Leu Gly Val Gly Phe Ala Ile Leu Tyr Thr Lys
     2150   2155   2160

CAC CGG AGG CTG CAG AGC AGC TTC ACC GCC TTC GCC AAC AGC CAC TAC   6611
His Arg Arg Leu Gln Ser Ser Phe Thr Ala Phe Ala Asn Ser His Tyr

```
              2165                2170                2175
AGC TCC AGG CTG GGG TCC GCA ATC TTC TCC TCT GGG GAT GAC CTG GGG    6659
Ser Ser Arg Leu Gly Ser Ala Ile Phe Ser Ser Gly Asp Asp Leu Gly
                2180                2185                2190
GAA GAT GAT GAA GAT GCC CCT ATG ATA ACT GGA TTT TCA GAT GAC GTC    6707
Glu Asp Asp Glu Asp Ala Pro Met Ile Thr Gly Phe Ser Asp Asp Val
        2195                2200                2205
CCC ATG GTG ATA GCC TGAAAGAGCT TTCCTCACTA GAAACCAAAT GGTGTAAATA    6762
Pro Met Val Ile Ala
2210
TTTTATTTGA TAAAGATAGT TGATGGTTTA TTTTAAAAGA TGCACTTTGA GTTGCAATAT    6822
GTTATTTTTA TATGGGCCAA A                                              6843
```

SEQUENCE LISTING

(1) GENERAL INFORMATION:

    (i) APPLICANT:
        (A) NAME: KOWA CO., LTD.
        (B) STREET: 6-29, Nishiki 3-chome, Naka-ku, Nagoya-shi,
        (C) CITY: Aichi
        (E) COUNTRY: Japan
        (F) POSTAL CODE (ZIP): none

    (ii) TITLE OF INVENTION: NOVEL LDL RECEPTOR ANALOG PROTEIN AND THE
        GENE CODING THEREFOR

  (iii) NUMBER OF SEQUENCES: 7

    (iv) COMPUTER READABLE FORM:
        (A) MEDIUM TYPE: Floppy disk
        (B) COMPUTER: IBM PC compatible
        (C) OPERATING SYSTEM: PC-DOS/MS-DOS
        (D) SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)

(2) INFORMATION FOR SEQ ID NO: 1:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 6639 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: double
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: cDNA to mRNA

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:

```
ATGGCGACAC GGAGCAGCAG GAGGGAGTCG CGACTCCCCT TCCTATTCAC CCTGGTCGCG      60
CTGCTGCCGC CCGGGGCTCT CTGCGAGGTG TGGACGCGGA CACTGCACGG CGGCCGCGCG     120
CCCTTACCCC AGGAGCGGGG CTTCCGCGTG GTGCAGGGCG ACCCGCGCGA GCTGCGGCTG     180
TGGGAGCGCG GGGATGCCAG GGGGGGCGAGC CGGGCGGACG AGAAGCCGCT CCGGAGGAGA     240
CGGAGCGCTG CCCTGCAGCC CGAGCCCATC AAGGTGTACG GACAGGTCAG CCTCAATGAT     300
TCCCACAATC AGATGGTGGT GCACTGGGCC GGAGAGAAAA GCAACGTGAT CGTGGCCTTG     360
GCCCGGGACA GCCTGGCGTT GGCCAGGCCC AGGAGCAGTG ATGTGTACGT GTCTTATGAC     420
TATGGAAAAT CATTCAATAA GATTTCAGAG AAAATTGAACT TCGGCGCGGG AAATAACACA     480
GAGGCTGTGG TGGCCCAGTT CTACCACAGC CCTGCGGACA ACAAACGGTA CATCTTCGCA     540
GATGCCTACG CCCAGTATCT CTGGATCACG TTTGACTTCT GCAACACCAT CCATGGCTTT     600
TCCATCCCGT TCCGGGCAGC TGATCTCCTA CTCCACAGTA AGGCCTCCAA CCTTCTCCTG     660
GGCTTCGACA GGTCTCACCC CAACAAGCAG CTGTGGAAGT CGGATGATTT TGGCCAGACC     720
TGGATCATGA TTCAAGAACA CGTGAAGTCC TTTTCTTGGG GAATTGATCC CTATGACAAA     780
```

CCAAACACCA	TCTACATCGA	ACGGCACGAA	CCTTCTGGCT	ACTCCACGGT	TTTCCGAAGT	840
ACAGACTTCT	TCCAGTCCCG	GGAAAACCAG	GAAGTGATCT	TGGAGGAAGT	GAGAGACTTT	900
CAGCTTCGGG	ACAAGTACAT	GTTTGCTACA	AAGGTGGTGC	ATCTCTTGGG	CAGTCCACTG	960
CAGTCTTCTG	TCCAGCTCTG	GGTCTCCTTT	GGCCGGAAGC	CCATGCGGGC	CGCCCAGTTT	1020
GTTACAAGAC	ATCCTATCAA	CGAATATTAC	ATCGCGGATG	CCTCGGAGGA	CCAGGTGTTT	1080
GTGTGTGTCA	GTCACAGCAA	CAACCGCACC	AACCTCTACA	TCTCGGAGGC	AGAGGGCTTG	1140
AAGTTCTCTC	TGTCCCTGGA	GAACGTGCTC	TACTACACCC	CGGGAGGGGC	CGGCAGTGAC	1200
ACCTTGGTGA	GGTACTTTGC	AAATGAACCG	TTTGCTGACT	TCCATCGTGT	GGAAGGGTTG	1260
CAGGGAGTCT	ACATTGCTAC	TCTGATTAAT	GGTTCTATGA	ATGAGGAGAA	CATGAGATCT	1320
GTCATCACCT	TTGACAAAGG	GGGCACCTGG	GAATTTCTGC	AGGCTCCAGC	CTTCACGGGG	1380
TATGGAGAGA	AAATCAACTG	TGAGCTGTCC	GAGGGCTGTT	CCCTCCACCT	GGCCCAGCGC	1440
CTCAGCCAGC	TGCTCAACCT	CCAGCTCCGG	AGGATGCCCA	TCCTGTCCAA	GGAGTCGGCG	1500
CCTGGCCTCA	TCATTGCCAC	GGGCTCAGTG	GGAAAGAACT	TGGCTAGCAA	GACAAACGTG	1560
TACATCTCTA	GCAGTGCTGG	AGCCAGGTGG	CGAGAGGCAC	TTCCTGGACC	TCACTACTAT	1620
ACATGGGGAG	ACCATGGCGG	CATCATCATG	GCCATTGCCC	AAGGCATGGA	AACCAACGAA	1680
CTGAAGTACA	GTACCAACGA	AGGGGAGACC	TGGAAAGCCT	TCACCTTCTC	TGAGAAGCCC	1740
GTGTTTGTGT	ATGGGCTCCT	CACGGAACCC	GGCGAGAAGA	GCACGGTCTT	CACCATCTTT	1800
GGCTCCAACA	AGGAGAACGT	GCACAGCTGG	CTCATCCTCC	AGGTCAATGC	CACAGACGCC	1860
CTGGGGGTTC	CTTGCACAGA	GAACGACTAC	AAGCTCTGGT	CACCATCTGA	TGAGCGGGGG	1920
AATGAGTGTT	TGCTTGGACA	CAAGACTGTT	TTCAAACGGA	GGACCCCGCA	CGCCACATGC	1980
TTTAACGGAG	AAGACTTTGA	CAGGCCGGTG	GTTGTGTCCA	ACTGCTCCTG	CACCCGGGAG	2040
GACTATGAGT	GTGACTTTGG	CTTCCGGATG	AGTGAAGACT	TGGCATTAGA	GGTGTGTGTT	2100
CCAGATCCAG	GATTTTCTGG	AAAGTCCTCC	CCTCCAGTGC	CTTGTCCCGT	GGGCTCTACG	2160
TACAGGCGAT	CAAGAGGCTA	CCGGAAGATT	TCTGGGGACA	CCTGTAGTGG	AGGAGATGTT	2220
GAGGCACGGC	TAGAAGGAGA	GCTGGTCCCC	TGTCCCCTGG	CAGAAGAGAA	CGAGTTCATC	2280
CTGTACGCCA	CGCGCAAGTC	CATCCACCGC	TATGACCTGG	CTTCCGGAAC	CACGGAGCAG	2340
TTGCCCCTCA	CTGGGTTGCG	GGCAGCAGTG	GCCCTGGACT	TTGACTATGA	GCACAACTGC	2400
CTGTATTGGT	CTGACCTGGC	CTTGGACGTC	ATCCAGCGCC	TCTGTTTGAA	CGGGAGTACA	2460
GGACAAGAGG	TGATCATCAA	CTCTGACCTG	GAGACGGTAG	AAGCTTTGGC	TTTTGAACCC	2520
CTCAGCCAAT	TACTTTACTG	GGTGGACGCA	GGCTTTAAAA	AGATCGAGGT	AGCCAATCCA	2580
GATGGTGACT	TCCGACTCAC	CGTCGTCAAT	TCCTCGGTGC	TGGATCGGCC	CCGGGCCCTG	2640
GTCCTTGTGC	CCCAAGAAGG	GATCATGTTC	TGGACCGACT	GGGGAGACCT	GAAGCCTGGG	2700

```
ATTTATCGGA GCAACATGGA CGGATCTGCC GCCTATCGCC TCGTGTCGGA GGATGTGAAG    2760
TGGCCCAATG GCATTTCCGT GGACGATCAG TGGATCTACT GGACGGATGC CTACCTGGAC    2820
TGCATTGAGC GCATCACGTT CAGCGGCCAG CAGCGCTCCG TCATCCTGGA CAGACTCCCG    2880
CACCCCTATG CCATTGCTGT CTTTAAGAAT GAGATTTACT GGGATGACTG GTCACAGCTC    2940
AGCATATTCC GAGCTTCTAA GTACAGCGGG TCCCAGATGG AGATTCTGGC CAGCCAGCTC    3000
ACGGGGCTGA TGGACATGAA GATCTTCTAC AAGGGGAAGA ACACAGGAAG CAATGCGTGT    3060
GTACCCAGGC CGTGCAGCCT GCTGTGCCTG CCCAGAGCCA ACAACAGCAA AAGCTGCAGG    3120
TGTCCAGATG GCGTGGCCAG CAGTGTCCTC CCTTCCGGGG ACCTGATGTG TGACTGCCCT    3180
AAGGGCTACG AGCTGAAGAA CAACACGTGT GTCAAAGAAG AAGACACCTG TCTGCGCAAC    3240
CAGTACCGCT GCAGCAACGG GAACTGCATC AACAGCATCT GGTGGTGCGA TTTCGACAAC    3300
GACTGCGGAG ACATGAGCGA CGAGAAGAAC TGCCCTACCA CCATCTGCGA CCTGGACACC    3360
CAGTTCCGTT GCCAGGAGTC TGGGACGTGC ATCCCGCTCT CCTACAAATG TGACCTCGAG    3420
GATGACTGTG GGGACAACAG TGACGAAAGG CACTGTGAAA TGCACCAGTG CCGGAGCGAC    3480
GAATACAACT GCAGCTCGGG CATGTGCATC CGCTCCTCCT GGGTGTGCGA CGGGGACAAC    3540
GACTGCAGGG ACTGGTCCGA CGAGGCCAAC TGCACAGCCA TCTATCACAC CTGTGAGGCC    3600
TCCAACTTCC AGTGCCGCAA CGGGCACTGC ATCCCCCAGC GGTGGGCGTG TGACGGCGAC    3660
GCCGACTGCC AGGATGGCTC TGATGAGGAT CCAGCCAACT GTGAGAAGAA GTGCAACGGC    3720
TTCCGCTGCC CGAACGGCAC CTGCATTCCC TCCACCAAGC ACTGTGACGG CCTGCACGAT    3780
TGCTCGGACG GCTCCGACGA GCAGCACTGC GAGCCCCTGT GTACACGGTT CATGGACTTC    3840
GTGTGTAAGA ACCGCCAGCA GTGCCTCTTC CACTCCATGG TGTGCGATGG GATCATCCAG    3900
TGCCGTGACG GCTCCGACGA GGACCCAGCC TTTGCAGGAT GCTCCCGAGA CCCCGAGTTC    3960
CACAAGGTGT GCGATGAGTT CGGCTTCCAG TGTCAGAACG GCGTGTGCAT CAGCTTGATC    4020
TGGAAGTGCG ACGGGATGGA TGACTGCGGG GACTACTCCG ACGAGGCCAA CTGTGAAAAC    4080
CCCACAGAAG CCCCCAACTG CTCCCGCTAC TTCCAGTTCC GGTGTGACAA TGGCCACTGC    4140
ATCCCCAACA GGTGGAAGTG TGACAGGGAG AATGACTGTG GGGACTGGTC CGACGAGAAG    4200
GACTGTGGAG ATTCACATGT ACTTCCGTCT ACGACTCCTG CACCCTCCAC GTGTCTGCCC    4260
AATTACTACC GCTGCGGCGG GGGGGCCTGC GTGATAGACA CGTGGGTTTG TGACGGGTAC    4320
CGAGATTGCG CAGATGGATC CGACGAGGAA GCCTGCCCCT CGCTCCCCAA TGTCACTGCC    4380
ACCTCCTCCC CCTCCCAGCC TGGACGATGC GACCGATTTG AGTTTGAGTG CCACCAGCCA    4440
AAGAAGTGCA TCCCTAACTG GAGACGCTGT GACGGCCATC AGGATTGCCA GGATGGCCAG    4500
GACGAGGCCA ACTGCCCCAC TCACAGCACC TTGACCTGCA TGAGCTGGGA GTTCAAGTGT    4560
GAGGATGGCG AGGCCTGCAT CGTGCTGTCA GAACGCTGCG ACGGCTTCCT GGACTGCTCA    4620
```

```
GATGAGAGCG ACGAGAAGGC CTGCAGTGAT GAGTTAACTG TATACAAAGT ACAGAATCTT    4680

CAGTGGACAG CTGACTTCTC TGGGAATGTC ACTTTGACCT GGATGCGGCC CAAAAAAATG    4740

CCCTCTGCTG CTTGTGTATA CAACGTGTAC TATAGAGTTG TTGGAGAGAG CATATGGAAG    4800

ACTCTGGAGA CTCACAGCAA TAAGACAAAC ACTGTATTAA AAGTGTTGAA ACCAGATACC    4860

ACCTACCAGG TTAAAGTGCA GGTTCAGTGC CTGAGCAAGG TGCACAACAC CAATGACTTT    4920

GTGACCTTGA GAACTCCAGA GGGATTGCCA GACGCCCCTC AGAACCTCCA GCTGTCGCTC    4980

CACGGGGAAG AGGAAGGTGT GATTGTGGGC CACTGGAGCC CTCCCACCCA CACCCACGGC    5040

CTCATTCGCG AATACATTGT AGAGTATAGC AGGAGTGGTT CCAAGGTGTG GACTTCAGAA    5100

AGGGCTGCTA GTAACTTTAC AGAAATAAAG AACTTGTTGG TCAACACCCT GTACACCGTC    5160

AGAGTGGCTG CGGTGACGAG TCGTGGGATA GGAAACTGGA GCGATTCCAA ATCCATTACC    5220

ACCGTGAAAG GAAAAGCGAT CCCGCCACCA AATATCCACA TTGACAACTA CGATGAAAAT    5280

TCCCTGAGTT TTACCCTGAC CGTGGATGGG AACATCAAGG TGAATGGCTA TGTGGTGAAC    5340

CTTTTCTGGG CATTTGACAC CCACAAACAA GAGAAGAAAA CCATGAACTT CCAAGGGAGC    5400

TCAGTGTCCC ACAAAGTTGG CAATCTGACA GCACAGACGG CCTATGAGAT TTCCGCCTGG    5460

GCCAAGACTG ACTTGGGCGA TAGTCCTCTG TCATTTGAGC ATGTCACGAC CAGAGGGGTT    5520

CGCCCACCTG CTCCTAGCCT CAAGGCCAGG GCTATCAATC AGACTGCAGT GGAATGCACC    5580

TGGACAGGCC CCAGGAATGT GGTGTATGGC ATTTTCTATG CCACATCCTT CCTGGACCTC    5640

TACCGCAACC CAAGCAGCCT GACCACGCCG CTGCACAACG CAACCGTGCT CGTCGGTAAG    5700

GATGAGCAGT ATCTGTTTCT GGTCCGGGTG GTGATGCCCT ACCAAGGGCC GTCCTCGGAC    5760

TACGTGGTCG TGAAGATGAT CCCGGACAGC AGGCTTCCTC CCCGGCACCT GCATGCCGTT    5820

CACACCGGCA AGACCTCGGC CGTCATCAAG TGGGAGTCGC CTACGACTC TCCTGACCAG    5880

GACCTGTTCT ATGCGATCGC AGTTAAAGAT CTGATACGAA AGACGGACCG GAGCTACAAA    5940

GTCAAGTCCC GCAACAGCAC CGTGGAGTAC ACCCTGAGCA AGCTGGAGCC CGGAGGGAAA    6000

TACCACGTCA TTGTGCAGCT GGGGAACATG AGCAAAGATG CCAGTGTGAA GATCACCACC    6060

GTTTCGTTAT CGGCACCCGA TGCCTTAAAA ATCATAACAG AAAATGACCA CGTCCTTCTC    6120

TTCTGGAAAA GTCTAGCTCT AAAGGAAAAG TATTTTAACG AAAGCAGGGG CTACGAGATA    6180

CACATGTTTG ATAGCGCCAT GAATATCACC GCATACCTTG GAATACTAC TGACAATTTC    6240

TTTAAAATTT CCAACCTGAA GATGGGTCAC AATTACACAT TCACGGTCCA GGCACGATGC    6300

CTTTTGGGCA GCCAGATCTG CGGGGAGCCT GCCGTGCTAC TGTATGATGA GCTGGGGTCT    6360

GGTGGCGATG CGTCGGCGAT GCAGGCTGCC AGGTCTACTG ATGTCGCCGC CGTGGTGGTG    6420

CCCATCCTGT TTCTGATACT GCTGAGCCTG GGGGTCGGGT TTGCCATCCT GTACACGAAG    6480

CATCGGAGGC TGCAGAGCAG CTTCACCGCC TTCGCCAACA GCCACTACAG CTCCAGACTC    6540
```

71

GGCTCCGCCA TCTTCTCCTC TGGGGATGAC TTGGGGGAGG ATGATGAAGA TGCTCCTATG   6600

ATCACTGGAT TTTCGGACGA CGTCCCCATG GTGATAGCC   .   6639

(2) INFORMATION FOR SEQ ID NO: 2:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 2213 amino acids
        (B) TYPE: amino acid
        (C) STRANDEDNESS:
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: protein

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:

```
Met Ala Thr Arg Ser Ser Arg Arg Glu Ser Arg Leu Pro Phe Leu Phe
1               5               10              15

Thr Leu Val Ala Leu Leu Pro Pro Gly Ala Leu Cys Glu Val Trp Thr
        20              25              30

Arg Thr Leu His Gly Gly Arg Ala Pro Leu Pro Gln Glu Arg Gly Phe
        35              40              45

Arg Val Val Gln Gly Asp Pro Arg Glu Leu Arg Leu Trp Glu Arg Gly
        50              55              60

Asp Ala Arg Gly Ala Ser Arg Ala Asp Glu Lys Pro Leu Arg Arg Arg
65              70              75              80

Arg Ser Ala Ala Leu Gln Pro Glu Pro Ile Lys Val Tyr Gly Gln Val
                85              90              95

Ser Leu Asn Asp Ser His Asn Gln Met Val Val His Trp Ala Gly Glu
            100             105             110

Lys Ser Asn Val Ile Val Ala Leu Ala Arg Asp Ser Leu Ala Leu Ala
            115             120             125

Arg Pro Arg Ser Ser Asp Val Tyr Val Ser Tyr Asp Tyr Gly Lys Ser
    130             135             140

Phe Asn Lys Ile Ser Glu Lys Leu Asn Phe Gly Ala Gly Asn Asn Thr
145             150             155             160

Glu Ala Val Val Ala Gln Phe Tyr His Ser Pro Ala Asp Asn Lys Arg
            165             170             175

Tyr Ile Phe Ala Asp Ala Tyr Ala Gln Tyr Leu Trp Ile Thr Phe Asp
            180             185             190

Phe Cys Asn Thr Ile His Gly Phe Ser Ile Pro Phe Arg Ala Ala Asp
            195             200             205

Leu Leu Leu His Ser Lys Ala Ser Asn Leu Leu Leu Gly Phe Asp Arg
    210             215             220

Ser His Pro Asn Lys Gln Leu Trp Lys Ser Asp Asp Phe Gly Gln Thr
225             230             235             240
```

```
Trp Ile Met Ile Gln Glu His Val Lys Ser Phe Ser Trp Gly Ile Asp
            245             250             255

Pro Tyr Asp Lys Pro Asn Thr Ile Tyr Ile Glu Arg His Glu Pro Ser
            260             265             270

Gly Tyr Ser Thr Val Phe Arg Ser Thr Asp Phe Phe Gln Ser Arg Glu
            275             280             285

Asn Gln Glu Val Ile Leu Glu Glu Val Arg Asp Phe Gln Leu Arg Asp
    290             295             300

Lys Tyr Met Phe Ala Thr Lys Val Val His Leu Leu Gly Ser Pro Leu
305             310             315             320

Gln Ser Ser Val Gln Leu Trp Val Ser Phe Gly Arg Lys Pro Met Arg
            325             330             335

Ala Ala Gln Phe Val Thr Arg His Pro Ile Asn Glu Tyr Tyr Ile Ala
            340             345             350

Asp Ala Ser Glu Asp Gln Val Phe Val Cys Val Ser His Ser Asn Asn
            355             360             365

Arg Thr Asn Leu Tyr Ile Ser Glu Ala Glu Gly Leu Lys Phe Ser Leu
    370             375             380

Ser Leu Glu Asn Val Leu Tyr Tyr Thr Pro Gly Gly Ala Gly Ser Asp
385             390             395             400

Thr Leu Val Arg Tyr Phe Ala Asn Glu Pro Phe Ala Asp Phe His Arg
            405             410             415

Val Glu Gly Leu Gln Gly Val Tyr Ile Ala Thr Leu Ile Asn Gly Ser
            420             425             430

Met Asn Glu Glu Asn Met Arg Ser Val Ile Thr Phe Asp Lys Gly Gly
    435             440             445

Thr Trp Glu Phe Leu Gln Ala Pro Ala Phe Thr Gly Tyr Gly Glu Lys
    450             455             460

Ile Asn Cys Glu Leu Ser Glu Gly Cys Ser Leu His Leu Ala Gln Arg
465             470             475             480

Leu Ser Gln Leu Leu Asn Leu Gln Leu Arg Arg Met Pro Ile Leu Ser
            485             490             495

Lys Glu Ser Ala Pro Gly Leu Ile Ile Ala Thr Gly Ser Val Gly Lys
            500             505             510

Asn Leu Ala Ser Lys Thr Asn Val Tyr Ile Ser Ser Ser Ala Gly Ala
            515             520             525

Arg Trp Arg Glu Ala Leu Pro Gly Pro His Tyr Tyr Thr Trp Gly Asp
            530             535             540

His Gly Gly Ile Ile Met Ala Ile Ala Gln Gly Met Glu Thr Asn Glu
545             550             555             560

Leu Lys Tyr Ser Thr Asn Glu Gly Glu Thr Trp Lys Ala Phe Thr Phe
            565             570             575
```

```
         Ser Glu Lys Pro Val Phe Val Tyr Gly Leu Leu Thr Glu Pro Gly Glu
                     580             585             590

    5    Lys Ser Thr Val Phe Thr Ile Phe Gly Ser Asn Lys Glu Asn Val His
                     595             600             605

         Ser Trp Leu Ile Leu Gln Val Asn Ala Thr Asp Ala Leu Gly Val Pro
             610             615             620

   10    Cys Thr Glu Asn Asp Tyr Lys Leu Trp Ser Pro Ser Asp Glu Arg Gly
         625             630             635             640

         Asn Glu Cys Leu Leu Gly His Lys Thr Val Phe Lys Arg Arg Thr Pro
                     645             650             655

   15    His Ala Thr Cys Phe Asn Gly Glu Asp Phe Asp Arg Pro Val Val Val
                     660             665             670

         Ser Asn Cys Ser Cys Thr Arg Glu Asp Tyr Glu Cys Asp Phe Gly Phe
                     675             680             685

         Arg Met Ser Glu Asp Leu Ala Leu Glu Val Cys Val Pro Asp Pro Gly
   20                690             695             700

         Phe Ser Gly Lys Ser Ser Pro Pro Val Pro Cys Pro Val Gly Ser Thr
         705             710             715             720

   25    Tyr Arg Arg Ser Arg Gly Tyr Arg Lys Ile Ser Gly Asp Thr Cys Ser
                         725             730             735

         Gly Gly Asp Val Glu Ala Arg Leu Glu Gly Glu Leu Val Pro Cys Pro
                     740             745             750

         Leu Ala Glu Glu Asn Glu Phe Ile Leu Tyr Ala Thr Arg Lys Ser Ile
   30                755             760             765

         His Arg Tyr Asp Leu Ala Ser Gly Thr Thr Glu Gln Leu Pro Leu Thr
                 770             775             780

         Gly Leu Arg Ala Ala Val Ala Leu Asp Phe Asp Tyr Glu His Asn Cys
         785             790             795             800

   35    Leu Tyr Trp Ser Asp Leu Ala Leu Asp Val Ile Gln Arg Leu Cys Leu
                         805             810             815

         Asn Gly Ser Thr Gly Gln Glu Val Ile Ile Asn Ser Asp Leu Glu Thr
                     820             825             830

   40    Val Glu Ala Leu Ala Phe Glu Pro Leu Ser Gln Leu Leu Tyr Trp Val
                     835             840             845

         Asp Ala Gly Phe Lys Lys Ile Glu Val Ala Asn Pro Asp Gly Asp Phe
             850             855             860

   45    Arg Leu Thr Val Val Asn Ser Ser Val Leu Asp Arg Pro Arg Ala Leu
         865             870             875             880

         Val Leu Val Pro Gln Glu Gly Ile Met Phe Trp Thr Asp Trp Gly Asp
                         885             890             895

   50    Leu Lys Pro Gly Ile Tyr Arg Ser Asn Met Asp Gly Ser Ala Ala Tyr
                     900             905             910

         Arg Leu Val Ser Glu Asp Val Lys Trp Pro Asn Gly Ile Ser Val Asp

   55
```

```
                   915                   920                   925
        Asp Gln Trp Ile Tyr Trp Thr Asp Ala Tyr Leu Asp Cys Ile Glu Arg
                930              935              940
        Ile Thr Phe Ser Gly Gln Gln Arg Ser Val Ile Leu Asp Arg Leu Pro
        945              950              955              960
        His Pro Tyr Ala Ile Ala Val Phe Lys Asn Glu Ile Tyr Trp Asp Asp
                        965              970              975
        Trp Ser Gln Leu Ser Ile Phe Arg Ala Ser Lys Tyr Ser Gly Ser Gln
                    980              985              990
        Met Glu Ile Leu Ala Ser Gln Leu Thr Gly Leu Met Asp Met Lys Ile
                    995              1000             1005
        Phe Tyr Lys Gly Lys Asn Thr Gly Ser Asn Ala Cys Val Pro Arg Pro
                1010             1015             1020
        Cys Ser Leu Leu Cys Leu Pro Arg Ala Asn Asn Ser Lys Ser Cys Arg
        1025             1030             1035             1040
        Cys Pro Asp Gly Val Ala Ser Ser Val Leu Pro Ser Gly Asp Leu Met
                    1045             1050             1055
        Cys Asp Cys Pro Lys Gly Tyr Glu Leu Lys Asn Asn Thr Cys Val Lys
                1060             1065             1070
        Glu Glu Asp Thr Cys Leu Arg Asn Gln Tyr Arg Cys Ser Asn Gly Asn
                1075             1080             1085
        Cys Ile Asn Ser Ile Trp Trp Cys Asp Phe Asp Asn Asp Cys Gly Asp
                1090             1095             1100
        Met Ser Asp Glu Lys Asn Cys Pro Thr Thr Ile Cys Asp Leu Asp Thr
        1105             1110             1115             1120
        Gln Phe Arg Cys Gln Glu Ser Gly Thr Cys Ile Pro Leu Ser Tyr Lys
                    1125             1130             1135
        Cys Asp Leu Glu Asp Asp Cys Gly Asp Asn Ser Asp Glu Arg His Cys
                1140             1145             1150
        Glu Met His Gln Cys Arg Ser Asp Glu Tyr Asn Cys Ser Ser Gly Met
                1155             1160             1165
        Cys Ile Arg Ser Ser Trp Val Cys Asp Gly Asp Asn Asp Cys Arg Asp
                1170             1175             1180
        Trp Ser Asp Glu Ala Asn Cys Thr Ala Ile Tyr His Thr Cys Glu Ala
        1185             1190             1195             1200
        Ser Asn Phe Gln Cys Arg Asn Gly His Cys Ile Pro Gln Arg Trp Ala
                    1205             1210             1215
        Cys Asp Gly Asp Ala Asp Cys Gln Asp Gly Ser Asp Glu Asp Pro Ala
                    1220             1225             1230
        Asn Cys Glu Lys Lys Cys Asn Gly Phe Arg Cys Pro Asn Gly Thr Cys
                1235             1240             1245
        Ile Pro Ser Thr Lys His Cys Asp Gly Leu His Asp Cys Ser Asp Gly
                1250             1255             1260
```

```
Ser Asp Glu Gln His Cys Glu Pro Leu Cys Thr Arg Phe Met Asp Phe
1265              1270          1275          1280

Val Cys Lys Asn Arg Gln Gln Cys Leu Phe His Ser Met Val Cys Asp
                1285              1290              1295

Gly Ile Ile Gln Cys Arg Asp Gly Ser Asp Glu Asp Pro Ala Phe Ala
            1300              1305              1310

Gly Cys Ser Arg Asp Pro Glu Phe His Lys Val Cys Asp Glu Phe Gly
            1315              1320              1325

Phe Gln Cys Gln Asn Gly Val Cys Ile Ser Leu Ile Trp Lys Cys Asp
        1330              1335              1340

Gly Met Asp Asp Cys Gly Asp Tyr Ser Asp Glu Ala Asn Cys Glu Asn
1345              1350              1355              1360

Pro Thr Glu Ala Pro Asn Cys Ser Arg Tyr Phe Gln Phe Arg Cys Asp
                1365              1370              1375

Asn Gly His Cys Ile Pro Asn Arg Trp Lys Cys Asp Arg Glu Asn Asp
            1380              1385              1390

Cys Gly Asp Trp Ser Asp Glu Lys Asp Cys Gly Asp Ser His Val Leu
        1395              1400              1405

Pro Ser Thr Thr Pro Ala Pro Ser Thr Cys Leu Pro Asn Tyr Tyr Arg
    1410              1415              1420

Cys Gly Gly Gly Ala Cys Val Ile Asp Thr Trp Val Cys Asp Gly Tyr
1425              1430              1435              1440

Arg Asp Cys Ala Asp Gly Ser Asp Glu Glu Ala Cys Pro Ser Leu Pro
            1445              1450              1455

Asn Val Thr Ala Thr Ser Ser Pro Ser Gln Pro Gly Arg Cys Asp Arg
            1460              1465              1470

Phe Glu Phe Glu Cys His Gln Pro Lys Lys Cys Ile Pro Asn Trp Arg
        1475              1480              1485

Arg Cys Asp Gly His Gln Asp Cys Gln Asp Gly Gln Asp Glu Ala Asn
    1490              1495              1500

Cys Pro Thr His Ser Thr Leu Thr Cys Met Ser Trp Glu Phe Lys Cys
1505              1510              1515              1520

Glu Asp Gly Glu Ala Cys Ile Val Leu Ser Glu Arg Cys Asp Gly Phe
                1525              1530              1535

Leu Asp Cys Ser Asp Glu Ser Asp Glu Lys Ala Cys Ser Asp Glu Leu
        1540              1545              1550

Thr Val Tyr Lys Val Gln Asn Leu Gln Trp Thr Ala Asp Phe Ser Gly
        1555              1560              1565

Asn Val Thr Leu Thr Trp Met Arg Pro Lys Lys Met Pro Ser Ala Ala
    1570              1575              1580

Cys Val Tyr Asn Val Tyr Tyr Arg Val Val Gly Glu Ser Ile Trp Lys
1585              1590              1595              1600
```

Thr Leu Glu Thr His Ser Asn Lys Thr Asn Thr Val Leu Lys Val Leu
                1605   1610   1615

Lys Pro Asp Thr Thr Tyr Gln Val Lys Val Gln Val Gln Cys Leu Ser
                1620   1625   1630

Lys Val His Asn Thr Asn Asp Phe Val Thr Leu Arg Thr Pro Glu Gly
            1635   1640   1645

Leu Pro Asp Ala Pro Gln Asn Leu Gln Leu Ser Leu His Gly Glu Glu
            1650   1655   1660

Glu Gly Val Ile Val Gly His Trp Ser Pro Pro Thr His Thr His Gly
1665   1670   1675   1680

Leu Ile Arg Glu Tyr Ile Val Glu Tyr Ser Arg Ser Gly Ser Lys Val
                1685   1690   1695

Trp Thr Ser Glu Arg Ala Ala Ser Asn Phe Thr Glu Ile Lys Asn Leu
                1700   1705   1710

Leu Val Asn Thr Leu Tyr Thr Val Arg Val Ala Ala Val Thr Ser Arg
            1715   1720   1725

Gly Ile Gly Asn Trp Ser Asp Ser Lys Ser Ile Thr Thr Val Lys Gly
    1730   1735   1740

Lys Ala Ile Pro Pro Pro Asn Ile His Ile Asp Asn Tyr Asp Glu Asn
1745   1750   1755   1760

Ser Leu Ser Phe Thr Leu Thr Val Asp Gly Asn Ile Lys Val Asn Gly
                1765   1770   1775

Tyr Val Val Asn Leu Phe Trp Ala Phe Asp Thr His Lys Gln Glu Lys
                1780   1785   1790

Lys Thr Met Asn Phe Gln Gly Ser Ser Val Ser His Lys Val Gly Asn
                1795   1800   1805

Leu Thr Ala Gln Thr Ala Tyr Glu Ile Ser Ala Trp Ala Lys Thr Asp
            1810   1815   1820

Leu Gly Asp Ser Pro Leu Ser Phe Glu His Val Thr Thr Arg Gly Val
1825   1830   1835   1840

Arg Pro Pro Ala Pro Ser Leu Lys Ala Arg Ala Ile Asn Gln Thr Ala
                1845   1850   1855

Val Glu Cys Thr Trp Thr Gly Pro Arg Asn Val Val Tyr Gly Ile Phe
                1860   1865   1870

Tyr Ala Thr Ser Phe Leu Asp Leu Tyr Arg Asn Pro Ser Ser Leu Thr
            1875   1880   1885

Thr Pro Leu His Asn Ala Thr Val Leu Val Gly Lys Asp Glu Gln Tyr
            1890   1895   1900

Leu Phe Leu Val Arg Val Val Met Pro Tyr Gln Gly Pro Ser Ser Asp
1905   1910   1915   1920

Tyr Val Val Val Lys Met Ile Pro Asp Ser Arg Leu Pro Pro Arg His
                1925   1930   1935

Leu His Ala Val His Thr Gly Lys Thr Ser Ala Val Ile Lys Trp Glu

```
                    1940                   1945                   1950
        Ser Pro Tyr Asp Ser Pro Asp Gln Asp Leu Phe Tyr Ala Ile Ala Val
                    1955                   1960               1965

        Lys Asp Leu Ile Arg Lys Thr Asp Arg Ser Tyr Lys Val Lys Ser Arg
                    1970                   1975               1980

        Asn Ser Thr Val Glu Tyr Thr Leu Ser Lys Leu Glu Pro Gly Gly Lys
        1985                    1990                   1995               2000

        Tyr His Val Ile Val Gln Leu Gly Asn Met Ser Lys Asp Ala Ser Val
                    2005                   2010                   2015

        Lys Ile Thr Thr Val Ser Leu Ser Ala Pro Asp Ala Leu Lys Ile Ile
                    2020                   2025                   2030

        Thr Glu Asn Asp His Val Leu Leu Phe Trp Lys Ser Leu Ala Leu Lys
                    2035                   2040                   2045

        Glu Lys Tyr Phe Asn Glu Ser Arg Gly Tyr Glu Ile His Met Phe Asp
                    2050                   2055                   2060

        Ser Ala Met Asn Ile Thr Ala Tyr Leu Gly Asn Thr Thr Asp Asn Phe
        2065                    2070                   2075               2080

        Phe Lys Ile Ser Asn Leu Lys Met Gly His Asn Tyr Thr Phe Thr Val
                    2085                   2090                   2095

        Gln Ala Arg Cys Leu Leu Gly Ser Gln Ile Cys Gly Glu Pro Ala Val
                    2100                   2105                   2110

        Leu Leu Tyr Asp Glu Leu Gly Ser Gly Gly Asp Ala Ser Ala Met Gln
                    2115                   2120                   2125

        Ala Ala Arg Ser Thr Asp Val Ala Ala Val Val Val Pro Ile Leu Phe
                    2130                   2135                   2140

        Leu Ile Leu Leu Ser Leu Gly Val Gly Phe Ala Ile Leu Tyr Thr Lys
        2145                    2150                   2155               2160

        His Arg Arg Leu Gln Ser Ser Phe Thr Ala Phe Ala Asn Ser His Tyr
                    2165                   2170                   2175

        Ser Ser Arg Leu Gly Ser Ala Ile Phe Ser Ser Gly Asp Asp Leu Gly
                    2180                   2185                   2190

        Glu Asp Asp Glu Asp Ala Pro Met Ile Thr Gly Phe Ser Asp Asp Val
                    2195                   2200                   2205

        Pro Met Val Ile Ala
                    2210
```

(2) INFORMATION FOR SEQ ID NO: 3:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 6961 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: double
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: cDNA to mRNA

```
      (ix) FEATURE:
           (A) NAME/KEY: sig peptide
           (B) LOCATION:178..261

      (ix) FEATURE:
           (A) NAME/KEY: mat peptide
           (B) LOCATION:262..6816


      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:

CCGCGAGCCG CACACGTGAC GGCGCCGCGC CGCGCCGCGC CGCGCCGAGC GGGACCCAGC   60
GGCTGCCCGG AGCCCCGGGA GCGGCGCGCG CGCGGCCCCG GCCCCGCCGC TCGGCCGGCG  120
GCGCGCTGCA CATTCTCTCC TGGCGGCGGC GCCACCTGCA GCCGCGTTCG CCCGAACATG  180
                                                                 Met
                                                                  1
GCG ACA CGG AGC AGC AGG AGG GAG TCG CGA CTC CCC TTC CTA TTC ACC   228
Ala Thr Arg Ser Ser Arg Arg Glu Ser Arg Leu Pro Phe Leu Phe Thr
              5                  10                  15
CTG GTC GCG CTG CTG CCG CCC GGG GCT CTC TGC GAG GTG TGG ACG CGG   276
Leu Val Ala Leu Leu Pro Pro Gly Ala Leu Cys Glu Val Trp Thr Arg
              20                  25                  30
ACA CTG CAC GGC GGC CGC GCG CCC TTA CCC CAG GAG CGG GGC TTC CGC   324
Thr Leu His Gly Gly Arg Ala Pro Leu Pro Gln Glu Arg Gly Phe Arg
          35              40                  45
GTG GTG CAG GGC GAC CCG CGC GAG CTG CGG CTG TGG GAG CGC GGG GAT   372
Val Val Gln Gly Asp Pro Arg Glu Leu Arg Leu Trp Glu Arg Gly Asp
50              55                  60                  65
GCC AGG GGG GCG AGC CGG GCG GAC GAG AAG CCG CTC CGG AGG AGA CGG   420
Ala Arg Gly Ala Ser Arg Ala Asp Glu Lys Pro Leu Arg Arg Arg Arg
              70                  75                  80
AGC GCT GCC CTG CAG CCC GAG CCC ATC AAG GTG TAC GGA CAG GTC AGC   468
Ser Ala Ala Leu Gln Pro Glu Pro Ile Lys Val Tyr Gly Gln Val Ser
              85                  90                  95
CTC AAT GAT TCC CAC AAT CAG ATG GTG GTG CAC TGG GCC GGA GAG AAA   516
Leu Asn Asp Ser His Asn Gln Met Val Val His Trp Ala Gly Glu Lys
          100                 105                 110
AGC AAC GTG ATC GTG GCC TTG GCC CGG GAC AGC CTG GCG TTG GCC AGG   564
Ser Asn Val Ile Val Ala Leu Ala Arg Asp Ser Leu Ala Leu Ala Arg
          115                 120                 125
CCC AGG AGC AGT GAT GTG TAC GTG TCT TAT GAC TAT GGA AAA TCA TTC   612
Pro Arg Ser Ser Asp Val Tyr Val Ser Tyr Asp Tyr Gly Lys Ser Phe
130                 135                 140                 145
AAT AAG ATT TCA GAG AAA TTG AAC TTC GGC GCG GGA AAT AAC ACA GAG   660
Asn Lys Ile Ser Glu Lys Leu Asn Phe Gly Ala Gly Asn Asn Thr Glu
                  150                 155                 160
GCT GTG GTG GCC CAG TTC TAC CAC AGC CCT GCG GAC AAC AAA CGG TAC   708
Ala Val Val Ala Gln Phe Tyr His Ser Pro Ala Asp Asn Lys Arg Tyr
              165                 170                 175
ATC TTC GCA GAT GCC TAC GCC CAG TAT CTC TGG ATC ACG TTT GAC TTC   756
Ile Phe Ala Asp Ala Tyr Ala Gln Tyr Leu Trp Ile Thr Phe Asp Phe
              180                 185                 190
TGC AAC ACC ATC CAT GGC TTT TCC ATC CCG TTC CGG GCA GCT GAT CTC   804
Cys Asn Thr Ile His Gly Phe Ser Ile Pro Phe Arg Ala Ala Asp Leu
          195                 200                 205
CTA CTC CAC AGT AAG GCC TCC AAC CTT CTC CTG GGC TTC GAC AGG TCT   852
Leu Leu His Ser Lys Ala Ser Asn Leu Leu Leu Gly Phe Asp Arg Ser
210                 215                 220                 225
CAC CCC AAC AAG CAG CTG TGG AAG TCG GAT GAT TTT GGC CAG ACC TGG   900
His Pro Asn Lys Gln Leu Trp Lys Ser Asp Asp Phe Gly Gln Thr Trp
                  230                 235                 240
ATC ATG ATT CAA GAA CAC GTG AAG TCC TTT TCT TGG GGA ATT GAT CCC   948
Ile Met Ile Gln Glu His Val Lys Ser Phe Ser Trp Gly Ile Asp Pro
              245                 250                 255
```

```
TAT GAC AAA CCA AAC ACC ATC TAC ATC GAA CGG CAC GAA CCT TCT GGC    996
Tyr Asp Lys Pro Asn Thr Ile Tyr Ile Glu Arg His Glu Pro Ser Gly
        260             265             270
TAC TCC ACG GTT TTC CGA AGT ACA GAC TTC TTC CAG TCC CGG GAA AAC   1044
Tyr Ser Thr Val Phe Arg Ser Thr Asp Phe Phe Gln Ser Arg Glu Asn
        275             280             285
CAG GAA GTG ATC TTG GAG GAA GTG AGA GAC TTT CAG CTT CGG GAC AAG   1092
Gln Glu Val Ile Leu Glu Glu Val Arg Asp Phe Gln Leu Arg Asp Lys
290             295             300             305
TAC ATG TTT GCT ACA AAG GTG GTG CAT CTC TTG GGC AGT CCA CTG CAG   1140
Tyr Met Phe Ala Thr Lys Val Val His Leu Leu Gly Ser Pro Leu Gln
        310             315             320
TCT TCT GTC CAG CTC TGG GTC TCC TTT GGC CGG AAG CCC ATG CGG GCC   1188
Ser Ser Val Gln Leu Trp Val Ser Phe Gly Arg Lys Pro Met Arg Ala
        325             330             335
GCC CAG TTT GTT ACA AGA CAT CCT ATC AAC GAA TAT TAC ATC GCG GAT   1236
Ala Gln Phe Val Thr Arg His Pro Ile Asn Glu Tyr Tyr Ile Ala Asp
        340             345             350
GCC TCG GAG GAC CAG GTG TTT GTG TGT GTC AGT CAC AGC AAC AAC CGC   1284
Ala Ser Glu Asp Gln Val Phe Val Cys Val Ser His Ser Asn Asn Arg
355             360             365
ACC AAC CTC TAC ATC TCG GAG GCA GAG GGC TTG AAG TTC TCT CTG TCC   1332
Thr Asn Leu Tyr Ile Ser Glu Ala Glu Gly Leu Lys Phe Ser Leu Ser
370             375             380             385
CTG GAG AAC GTG CTC TAC TAC ACC CCG GGA GGG GCC GGC AGT GAC ACC   1380
Leu Glu Asn Val Leu Tyr Tyr Thr Pro Gly Gly Ala Gly Ser Asp Thr
            390             395             400
TTG GTG AGG TAC TTT GCA AAT GAA CCG TTT GCT GAC TTC CAT CGT GTG   1428
Leu Val Arg Tyr Phe Ala Asn Glu Pro Phe Ala Asp Phe His Arg Val
            405             410             415
GAA GGG TTG CAG GGA GTC TAC ATT GCT ACT CTG ATT AAT GGT TCT ATG   1476
Glu Gly Leu Gln Gly Val Tyr Ile Ala Thr Leu Ile Asn Gly Ser Met
        420             425             430
AAT GAG GAG AAC ATG AGA TCT GTC ATC ACC TTT GAC AAA GGG GGC ACC   1524
Asn Glu Glu Asn Met Arg Ser Val Ile Thr Phe Asp Lys Gly Gly Thr
    435             440             445
TGG GAA TTT CTG CAG GCT CCA GCC TTC ACG GGG TAT GGA GAG AAA ATC   1572
Trp Glu Phe Leu Gln Ala Pro Ala Phe Thr Gly Tyr Gly Glu Lys Ile
450             455             460             465
AAC TGT GAG CTG TCC GAG GGC TGT TCC CTC CAC CTG GCC CAG CGC CTC   1620
Asn Cys Glu Leu Ser Glu Gly Cys Ser Leu His Leu Ala Gln Arg Leu
            470             475             480
AGC CAG CTG CTC AAC CTC CAG CTC CGG AGG ATG CCC ATC CTG TCC AAG   1668
Ser Gln Leu Leu Asn Leu Gln Leu Arg Arg Met Pro Ile Leu Ser Lys
        485             490             495
GAG TCG GCG CCT GGC CTC ATC ATT GCC ACG GGC TCA GTG GGA AAG AAC   1716
Glu Ser Ala Pro Gly Leu Ile Ile Ala Thr Gly Ser Val Gly Lys Asn
500             505             510
TTG GCT AGC AAG ACA AAC GTG TAC ATC TCT AGC AGT GCT GGA GCC AGG   1764
Leu Ala Ser Lys Thr Asn Val Tyr Ile Ser Ser Ser Ala Gly Ala Arg
    515             520             525
TGG CGA GAG GCA CTT CCT GGA CCT CAC TAC TAT ACA TGG GGA GAC CAT   1812
Trp Arg Glu Ala Leu Pro Gly Pro His Tyr Tyr Thr Trp Gly Asp His
530             535             540             545
GGC GGC ATC ATC ATG GCC ATT GCC CAA GGC ATG GAA ACC AAC GAA CTG   1860
Gly Gly Ile Ile Met Ala Ile Ala Gln Gly Met Glu Thr Asn Glu Leu
            550             555             560
AAG TAC AGT ACC AAC GAA GGG GAG ACC TGG AAA GCC TTC ACC TTC TCT   1908
Lys Tyr Ser Thr Asn Glu Gly Glu Thr Trp Lys Ala Phe Thr Phe Ser
        565             570             575
GAG AAG CCC GTG TTT GTG TAT GGG CTC CTC ACG GAA CCC GGC GAG AAG   1956
Glu Lys Pro Val Phe Val Tyr Gly Leu Leu Thr Glu Pro Gly Glu Lys
        580             585             590
AGC ACG GTC TTC ACC ATC TTT GGC TCC AAC AAG GAG AAC GTG CAC AGC   2004
```

```
       Ser Thr Val Phe Thr Ile Phe Gly Ser Asn Lys Glu Asn Val His Ser
           595             600             605
       TGG CTC ATC CTC CAG GTC AAT GCC ACA GAC GCC CTG GGG GTT CCT TGC      2052
 5     Trp Leu Ile Leu Gln Val Asn Ala Thr Asp Ala Leu Gly Val Pro Cys
       610             615             620             625
       ACA GAG AAC GAC TAC AAG CTC TGG TCA CCA TCT GAT GAG CGG GGG AAT      2100
       Thr Glu Asn Asp Tyr Lys Leu Trp Ser Pro Ser Asp Glu Arg Gly Asn
                       630             635             640
       GAG TGT TTG CTT GGA CAC AAG ACT GTT TTC AAA CGG AGG ACC CCG CAC      2148
10     Glu Cys Leu Leu Gly His Lys Thr Val Phe Lys Arg Arg Thr Pro His
                   645             650             655
       GCC ACA TGC TTT AAC GGA GAA GAC TTT GAC AGG CCG GTG GTT GTG TCC      2196
       Ala Thr Cys Phe Asn Gly Glu Asp Phe Asp Arg Pro Val Val Val Ser
               660             665             670
       AAC TGC TCC TGC ACC CGG GAG GAC TAT GAG TGT GAC TTT GGC TTC CGG      2244
15     Asn Cys Ser Cys Thr Arg Glu Asp Tyr Glu Cys Asp Phe Gly Phe Arg
               675             680             685
       ATG AGT GAA GAC TTG GCA TTA GAG GTG TGT GTT CCA GAT CCA GGA TTT      2292
       Met Ser Glu Asp Leu Ala Leu Glu Val Cys Val Pro Asp Pro Gly Phe
       690             695             700             705
       TCT GGA AAG TCC TCC CCT CCA GTG CCT TGT CCC GTG GGC TCT ACG TAC      2340
       Ser Gly Lys Ser Ser Pro Pro Val Pro Cys Pro Val Gly Ser Thr Tyr
20                     710             715             720
       AGG CGA TCA AGA GGC TAC CGG AAG ATT TCT GGG GAC ACC TGT AGT GGA      2388
       Arg Arg Ser Arg Gly Tyr Arg Lys Ile Ser Gly Asp Thr Cys Ser Gly
                       725             730             735
       GGA GAT GTT GAG GCA CGG CTA GAA GGA GAG CTG GTC CCC TGT CCC CTG      2436
       Gly Asp Val Glu Ala Arg Leu Glu Gly Glu Leu Val Pro Cys Pro Leu
                   740             745             750
25     GCA GAA GAG AAC GAG TTC ATC CTG TAC GCC ACG CGC AAG TCC ATC CAC      2484
       Ala Glu Glu Asn Glu Phe Ile Leu Tyr Ala Thr Arg Lys Ser Ile His
               755             760             765
       CGC TAT GAC CTG GCT TCC GGA ACC ACG GAG CAG TTG CCC CTC ACT GGG      2532
       Arg Tyr Asp Leu Ala Ser Gly Thr Thr Glu Gln Leu Pro Leu Thr Gly
       770             775             780             785
30     TTG CGG GCA GCA GTG GCC CTG GAC TTT GAC TAT GAG CAC AAC TGC CTG      2580
       Leu Arg Ala Ala Val Ala Leu Asp Phe Asp Tyr Glu His Asn Cys Leu
                       790             795             800
       TAT TGG TCT GAC CTG GCC TTG GAC GTC ATC CAG CGC CTC TGT TTG AAC      2628
       Tyr Trp Ser Asp Leu Ala Leu Asp Val Ile Gln Arg Leu Cys Leu Asn
                   805             810             815
35     GGG AGT ACA GGA CAA GAG GTG ATC ATC AAC TCT GAC CTG GAG ACG GTA      2676
       Gly Ser Thr Gly Gln Glu Val Ile Ile Asn Ser Asp Leu Glu Thr Val
               820             825             830
       GAA GCT TTG GCT TTT GAA CCC CTC AGC CAA TTA CTT TAC TGG GTG GAC      2724
       Glu Ala Leu Ala Phe Glu Pro Leu Ser Gln Leu Leu Tyr Trp Val Asp
               835             840             845
       GCA GGC TTT AAA AAG ATC GAG GTA GCC AAT CCA GAT GGT GAC TTC CGA      2772
40     Ala Gly Phe Lys Lys Ile Glu Val Ala Asn Pro Asp Gly Asp Phe Arg
       850             855             860             865
       CTC ACC GTC GTC AAT TCC TCG GTG CTG GAT CGG CCC CGG GCC CTG GTC      2820
       Leu Thr Val Val Asn Ser Ser Val Leu Asp Arg Pro Arg Ala Leu Val
                       870             875             880
       CTT GTG CCC CAA GAA GGG ATC ATG TTC TGG ACC GAC TGG GGA GAC CTG      2868
45     Leu Val Pro Gln Glu Gly Ile Met Phe Trp Thr Asp Trp Gly Asp Leu
                       885             890             895
       AAG CCT GGG ATT TAT CGG AGC AAC ATG GAC GGA TCT GCC GCC TAT CGC      2916
       Lys Pro Gly Ile Tyr Arg Ser Asn Met Asp Gly Ser Ala Ala Tyr Arg
                   900             905             910
       CTC GTG TCG GAG GAT GTG AAG TGG CCC AAT GGC ATT TCC GTG GAC GAT      2964
50     Leu Val Ser Glu Asp Val Lys Trp Pro Asn Gly Ile Ser Val Asp Asp
       915             920             925
       CAG TGG ATC TAC TGG ACG GAT GCC TAC CTG GAC TGC ATT GAG CGC ATC      3012
       Gln Trp Ile Tyr Trp Thr Asp Ala Tyr Leu Asp Cys Ile Glu Arg Ile
```

```
      930                    935                 940                   945
      ACG TTC AGC GGC CAG CAG CGC TCC GTC ATC CTG GAC ACA CTC CCG CAC    3060
      Thr Phe Ser Gly Gln Gln Arg Ser Val Ile Leu Asp Arg Leu Pro His
                        950                 955                 960
      CCC TAT GCC ATT GCT GTC TTT AAG AAT GAG ATT TAC TGG GAT GAC TGG    3108
      Pro Tyr Ala Ile Ala Val Phe Lys Asn Glu Ile Tyr Trp Asp Asp Trp
                        965                 970                 975
      TCA CAG CTC AGC ATA TTC CGA GCT TCT AAG TAC AGC GGG TCC CAG ATG    3156
      Ser Gln Leu Ser Ile Phe Arg Ala Ser Lys Tyr Ser Gly Ser Gln Met
                    980                 985                 990
      GAG ATT CTG GCC AGC CAG CTC ACG GGG CTG ATG GAC ATG AAG ATC TTC    3204
      Glu Ile Leu Ala Ser Gln Leu Thr Gly Leu Met Asp Met Lys Ile Phe
              995                 1000                1005
      TAC AAG GGG AAG AAC ACA GGA AGC AAT GCG TGT GTA CCC AGG CCG TGC    3252
      Tyr Lys Gly Lys Asn Thr Gly Ser Asn Ala Cys Val Pro Arg Pro Cys
      1010                1015                1020                1025
      AGC CTG CTG TGC CTG CCC AGA GCC AAC AAC AGC AAA AGC TGC AGG TGT    3300
      Ser Leu Leu Cys Leu Pro Arg Ala Asn Asn Ser Lys Ser Cys Arg Cys
                        1030                1035                1040
      CCA GAT GGC GTG GCC AGC AGT GTC CTC CCT TCC GGG GAC CTG ATG TGT    3348
      Pro Asp Gly Val Ala Ser Ser Val Leu Pro Ser Gly Asp Leu Met Cys
                    1045                1050                1055
      GAC TGC CCT AAG GGC TAC GAG CTG AAG AAC AAC ACG TGT GTC AAA GAA    3396
      Asp Cys Pro Lys Gly Tyr Glu Leu Lys Asn Asn Thr Cys Val Lys Glu
              1060                1065                1070
      GAA GAC ACC TGT CTG CGC AAC CAG TAC CGC TGC AGC AAC GGG AAC TGC    3444
      Glu Asp Thr Cys Leu Arg Asn Gln Tyr Arg Cys Ser Asn Gly Asn Cys
          1075                1080                1085
      ATC AAC AGC ATC TGG TGG TGC GAT TTC GAC AAC GAC TGC GGA GAC ATG    3492
      Ile Asn Ser Ile Trp Trp Cys Asp Phe Asp Asn Asp Cys Gly Asp Met
      1090                1095                1100                1105
      AGC GAC GAG AAG AAC TGC CCT ACC ACC ATC TGC GAC CTG GAC ACC CAG    3540
      Ser Asp Glu Lys Asn Cys Pro Thr Thr Ile Cys Asp Leu Asp Thr Gln
                        1110                1115                1120
      TTC CGT TGC CAG GAG TCT GGG ACG TGC ATC CCG CTC TCC TAC AAA TGT    3588
      Phe Arg Cys Gln Glu Ser Gly Thr Cys Ile Pro Leu Ser Tyr Lys Cys
                    1125                1130                1135
      GAC CTC GAG GAT GAC TGT GGG GAC AAC AGT GAC GAA AGG CAC TGT GAA    3636
      Asp Leu Glu Asp Asp Cys Gly Asp Asn Ser Asp Glu Arg His Cys Glu
              1140                1145                1150
      ATG CAC CAG TGC CGG AGC GAC GAA TAC AAC TGC AGC TCG GGC ATG TGC    3684
      Met His Gln Cys Arg Ser Asp Glu Tyr Asn Cys Ser Ser Gly Met Cys
          1155                1160                1165
      ATC CGC TCC TCC TGG GTG TGC GAC GGG GAC AAC GAC TGC AGG GAC TGG    3732
      Ile Arg Ser Ser Trp Val Cys Asp Gly Asp Asn Asp Cys Arg Asp Trp
      1170                1175                1180                1185
      TCC GAC GAG GCC AAC TGC ACA GCC ATC TAT CAC ACC TGT GAG GCC TCC    3780
      Ser Asp Glu Ala Asn Cys Thr Ala Ile Tyr His Thr Cys Glu Ala Ser
                        1190                1195                1200
      AAC TTC CAG TGC CGC AAC GGG CAC TGC ATC CCC CAG CGG TGG GCG TGT    3828
      Asn Phe Gln Cys Arg Asn Gly His Cys Ile Pro Gln Arg Trp Ala Cys
                    1205                1210                1215
      GAC GGC GAC GCC GAC TGC CAG GAT GGC TCT GAT GAG GAT CCA GCC AAC    3876
      Asp Gly Asp Ala Asp Cys Gln Asp Gly Ser Asp Glu Asp Pro Ala Asn
              1220                1225                1230
      TGT GAG AAG AAG TGC AAC GGC TTC CGC TGC CCG AAC GGC ACC TGC ATT    3924
      Cys Glu Lys Lys Cys Asn Gly Phe Arg Cys Pro Asn Gly Thr Cys Ile
          1235                1240                1245
      CCC TCC ACC AAG CAC TGT GAC GGC CTG CAC GAT TGC TCG GAC GGC TCC    3972
      Pro Ser Thr Lys His Cys Asp Gly Leu His Asp Cys Ser Asp Gly Ser
      1250                1255                1260                1265
      GAC GAG CAG CAC TGC GAG CCC CTG TGT ACA CGG TTC ATG GAC TTC GTG    4020
      Asp Glu Gln His Cys Glu Pro Leu Cys Thr Arg Phe Met Asp Phe Val
                    1270                1275                1280
```

```
TGT AAG AAC CGC CAG CAG TGC CTC TTC CAC TCC ATG CTG TGC GAT GGG    4063
Cys Lys Asn Arg Gln Gln Cys Leu Phe His Ser Met Val Cys Asp Gly
            1285                1290                1295
ATC ATC CAG TGC CGT GAC GGC TCC GAC GAG GAC CCA GCC TTT GCA GGA    4116
Ile Ile Gln Cys Arg Asp Gly Ser Asp Glu Asp Pro Ala Phe Ala Gly
            1300                1305                1310
TGC TCC CGA GAC CCC GAG TTC CAC AAG GTG TGC GAT GAG TTC GGC TTC    4164
Cys Ser Arg Asp Pro Glu Phe His Lys Val Cys Asp Glu Phe Gly Phe
        1315                1320                1325
CAG TGT CAG AAC GGC GTG TGC ATC AGC TTG ATC TGG AAG TGC GAC GGG    4212
Gln Cys Gln Asn Gly Val Cys Ile Ser Leu Ile Trp Lys Cys Asp Gly
1330                1335                1340                1345
ATG GAT GAC TGC GGG GAC TAC TCC GAC GAG GCC AAC TGT GAA AAC CCC    4260
Met Asp Asp Cys Gly Asp Tyr Ser Asp Glu Ala Asn Cys Glu Asn Pro
                1350                1355                1360
ACA GAA GCC CCC AAC TGC TCC CGC TAC TTC CAG TTC CGG TGT GAC AAT    4308
Thr Glu Ala Pro Asn Cys Ser Arg Tyr Phe Gln Phe Arg Cys Asp Asn
            1365                1370                1375
GGC CAC TGC ATC CCC AAC AGG TGG AAG TGT GAC AGG GAG AAT GAC TGT    4356
Gly His Cys Ile Pro Asn Arg Trp Lys Cys Asp Arg Glu Asn Asp Cys
        1380                1385                1390
GGG GAC TGG TCC GAC GAG AAG GAC TGT GGA GAT TCA CAT GTA CTT CCG    4404
Gly Asp Trp Ser Asp Glu Lys Asp Cys Gly Asp Ser His Val Leu Pro
        1395                1400                1405
TCT ACG ACT CCT GCA CCC TCC ACG TGT CTG CCC AAT TAC TAC CGC TGC    4452
Ser Thr Thr Pro Ala Pro Ser Thr Cys Leu Pro Asn Tyr Tyr Arg Cys
1410                1415                1420                1425
GGC GGG GGG GCC TGC GTG ATA GAC ACG TGG GTT TGT GAC GGG TAC CGA    4500
Gly Gly Gly Ala Cys Val Ile Asp Thr Trp Val Cys Asp Gly Tyr Arg
                1430                1435                1440
GAT TGC GCA GAT GGA TCC GAC GAG GAA GCC TGC CCC TCG CTC CCC AAT    4548
Asp Cys Ala Asp Gly Ser Asp Glu Glu Ala Cys Pro Ser Leu Pro Asn
                1445                1450                1455
GTC ACT GCC ACC TCC TCC CCC TCC CAG CCT GGA CGA TGC GAC CGA TTT    4596
Val Thr Ala Thr Ser Ser Pro Ser Gln Pro Gly Arg Cys Asp Arg Phe
            1460                1465                1470
GAG TTT GAG TGC CAC CAG CCA AAG AAG TGC ATC CCT AAC TGG AGA CGC    4644
Glu Phe Glu Cys His Gln Pro Lys Lys Cys Ile Pro Asn Trp Arg Arg
            1775                1480                1485
TGT GAC GGC CAT CAG GAT TGC CAG GAT GGC CAG GAC GAG GCC AAC TGC    4692
Cys Asp Gly His Gln Asp Cys Gln Asp Gly Gln Asp Glu Ala Asn Cys
1490                1495                1500                1505
CCC ACT CAC AGC ACC TTG ACC TGC ATG AGC TGG GAG TTC AAG TGT GAG    4740
Pro Thr His Ser Thr Leu Thr Cys Met Ser Trp Glu Phe Lys Cys Glu
                1510                1515                1520
GAT GGC GAG GCC TGC ATC GTG CTG TCA GAA CGC TGC GAC GGC TTC CTG    4788
Asp Gly Glu Ala Cys Ile Val Leu Ser Glu Arg Cys Asp Gly Phe Leu
            1525                1530                1535
GAC TGC TCA GAT GAG AGC GAC GAG AAG GCC TGC AGT GAT GAG TTA ACT    4836
Asp Cys Ser Asp Glu Ser Asp Glu Lys Ala Cys Ser Asp Glu Leu Thr
            1540                1545                1550
GTA TAC AAA GTA CAG AAT CTT CAG TGG ACA GCT GAC TTC TCT GGG AAT    4884
Val Tyr Lys Val Gln Asn Leu Gln Trp Thr Ala Asp Phe Ser Gly Asn
            1555                1560                1565
GTC ACT TTG ACC TGG ATG CGG CCC AAA AAA ATG CCC TCT GCT GCT TGT    4932
Val Thr Leu Thr Trp Met Arg Pro Lys Lys Met Pro Ser Ala Ala Cys
1570                1575                1580                1585
GTA TAC AAC GTG TAC TAT AGA GTT GTT GGA GAG AGC ATA TGG AAG ACT    4980
Val Tyr Asn Val Tyr Tyr Arg Val Val Gly Glu Ser Ile Trp Lys Thr
                1590                1595                1600
CTG GAG ACT CAC AGC AAT AAG ACA AAC ACT GTA TTA AAA GTG TTG AAA    5028
Leu Glu Thr His Ser Asn Lys Thr Asn Thr Val Leu Lys Val Leu Lys
        1605                1610                1615
CCA GAT ACC ACC TAC CAG GTT AAA GTG CAG GTT CAG TGC CTG AGC AAG    5076
```

83

```
    Pro Asp Thr Thr Tyr Gln Val Lys Val Gln Val Gln Cys Leu Ser Lys
            1620                1625                1630
    GTG CAC AAC ACC AAT GAC TTT GTG ACC TTG AGA ACT CCA GAG GGA TTG    5124
    Val His Asn Thr Asn Asp Phe Val Thr Leu Arg Thr Pro Glu Gly Leu
        1635                1640                1645
    CCA GAC GCC CCT CAG AAC CTC CAG CTG TCG CTC CAC GGG GAA GAG GAA    5172
    Pro Asp Ala Pro Gln Asn Leu Gln Leu Ser Leu His Gly Glu Glu Glu
    1650                1655                1660                1665
    GGT GTG ATT GTG GGC CAC TGG AGC CCT CCC ACC CAC ACC CAC GGC CTC    5220
    Gly Val Ile Val Gly His Trp Ser Pro Pro Thr His Thr His Gly Leu
                    1670                1675                1680
    ATT CGC GAA TAC ATT GTA GAG TAT AGC AGG AGT GGT TCC AAG GTG TGG    5268
    Ile Arg Glu Tyr Ile Val Glu Tyr Ser Arg Ser Gly Ser Lys Val Trp
                1685                1690                1695
    ACT TCA GAA AGG GCT GCT AGT AAC TTT ACA GAA ATA AAG AAC TTG TTG    5316
    Thr Ser Glu Arg Ala Ala Ser Asn Phe Thr Glu Ile Lys Asn Leu Leu
            1700                1705                1710
    GTC AAC ACC CTG TAC ACC GTC AGA GTG GCT GCG GTG ACG AGT CGT GGG    5364
    Val Asn Thr Leu Tyr Thr Val Arg Val Ala Ala Val Thr Ser Arg Gly
        1715                1720                1725
    ATA GGA AAC TGG AGC GAT TCC AAA TCC ATT ACC ACC GTG AAA GGA AAA    5412
    Ile Gly Asn Trp Ser Asp Ser Lys Ser Ile Thr Thr Val Lys Gly Lys
    1730                1735                1740                1745
    GCG ATC CCG CCA CCA AAT ATC CAC ATT GAC AAC TAC GAT GAA AAT TCC    5460
    Ala Ile Pro Pro Pro Asn Ile His Ile Asp Asn Tyr Asp Glu Asn Ser
                    1750                1755                1760
    CTG AGT TTT ACC CTG ACC GTG GAT GGG AAC ATC AAG GTG AAT GGC TAT    5508
    Leu Ser Phe Thr Leu Thr Val Asp Gly Asn Ile Lys Val Asn Gly Tyr
                1765                1770                1775
    GTG GTG AAC CTT TTC TGG GCA TTT GAC ACC CAC AAA CAA GAG AAG AAA    5556
    Val Val Asn Leu Phe Trp Ala Phe Asp Thr His Lys Gln Glu Lys Lys
            1780                1785                1790
    ACC ATG AAC TTC CAA GGG AGC TCA GTG TCC CAC AAA GTT GGC AAT CTG    5604
    Thr Met Asn Phe Gln Gly Ser Ser Val Ser His Lys Val Gly Asn Leu
        1795                1800                1805
    ACA GCA CAG ACG GCC TAT GAG ATT TCC GCC TGG GCC AAG ACT GAC TTG    5652
    Thr Ala Gln Thr Ala Tyr Glu Ile Ser Ala Trp Ala Lys Thr Asp Leu
    1810                1815                1820                1825
    GGC GAT AGT CCT CTG TCA TTT GAG CAT GTC ACG ACC AGA GGG GTT CGC    5700
    Gly Asp Ser Pro Leu Ser Phe Glu His Val Thr Thr Arg Gly Val Arg
                    1830                1835                1840
    CCA CCT GCT CCT AGC CTC AAG GCC AGG GCT ATC AAT CAG ACT GCA GTG    5748
    Pro Pro Ala Pro Ser Leu Lys Ala Arg Ala Ile Asn Gln Thr Ala Val
                1845                1850                1855
    GAA TGC ACC TGG ACA GGC CCC AGG AAT GTG GTG TAT GGC ATT TTC TAT    5796
    Glu Cys Thr Trp Thr Gly Pro Arg Asn Val Val Tyr Gly Ile Phe Tyr
            1860                1865                1870
    GCC ACA TCC TTC CTG GAC CTC TAC CGC AAC CCA AGC AGC CTG ACC ACG    5844
    Ala Thr Ser Phe Leu Asp Leu Tyr Arg Asn Pro Ser Ser Leu Thr Thr
        1875                1880                1885
    CCG CTG CAC AAC GCA ACC GTG CTC GTC GGT AAG GAT GAG CAG TAT CTG    5892
    Pro Leu His Asn Ala Thr Val Leu Val Gly Lys Asp Glu Gln Tyr Leu
    1890                1895                1900                1905
    TTT CTG GTC CGG GTG GTG ATG CCC TAC CAA GGG CCG TCC TCG GAC TAC    5940
    Phe Leu Val Arg Val Val Met Pro Tyr Gln Gly Pro Ser Ser Asp Tyr
                    1910                1915                1920
    GTG GTC GTG AAG ATG ATC CCG GAC AGC AGG CTT CCT CCC CGG CAC CTG    5988
    Val Val Val Lys Met Ile Pro Asp Ser Arg Leu Pro Pro Arg His Leu
                1925                1930                1935
    CAT GCC GTT CAC ACC GGC AAG ACC TCG GCC GTC ATC AAG TGG GAG TCG    6036
    His Ala Val His Thr Gly Lys Thr Ser Ala Val Ile Lys Trp Glu Ser
            1940                1945                1950
    CCC TAC GAC TCT CCT GAC CAG GAC CTG TTC TAT GCG ATC GCA GTT AAA    6084
    Pro Tyr Asp Ser Pro Asp Gln Asp Leu Phe Tyr Ala Ile Ala Val Lys
```

```
          1955                    1960                    1965
    GAT CTG ATA CGA AAG ACG GAC CGG AGC TAC AAA GTC AAG TCC CGC AAC   6132
    Asp Leu Ile Arg Lys Thr Asp Arg Ser Tyr Lys Val Lys Ser Arg Asn
    1970                    1975                    1980         .1985
    AGC ACC GTG GAG TAC ACC CTG AGC AAG CTG GAG CCC GGA GGG AAA TAC   6180
    Ser Thr Val Glu Tyr Thr Leu Ser Lys Leu Glu Pro Gly Gly Lys Tyr
                   1990                    1995                    2000
    CAC GTC ATT GTG CAG CTG GGG AAC ATG AGC AAA GAT GCC AGT GTG AAG   6228
    His Val Ile Val Gln Leu Gly Asn Met Ser Lys Asp Ala Ser Val Lys
                   2005                    2010                    2015
    ATC ACC ACC GTT TCG TTA TCG GCA CCC GAT GCC TTA AAA ATC ATA ACA   6276
    Ile Thr Thr Val Ser Leu Ser Ala Pro Asp Ala Leu Lys Ile Ile Thr
              2020                    2025                    2030
    GAA AAT GAC CAC GTC CTT CTC TTC TGG AAA AGT CTA GCT CTA AAG GAA   6324
    Glu Asn Asp His Val Leu Leu Phe Trp Lys Ser Leu Ala Leu Lys Glu
                   2035                    2040                    2045
    AAG TAT TTT AAC GAA AGC AGG GGC TAC GAG ATA CAC ATG TTT GAT AGC   6372
    Lys Tyr Phe Asn Glu Ser Arg Gly Tyr Glu Ile His Met Phe Asp Ser
    2050                    2055                    2060         2065
    GCC ATG AAT ATC ACC GCA TAC CTT GGG AAT ACT ACT GAC AAT TTC TTT   6420
    Ala Met Asn Ile Thr Ala Tyr Leu Gly Asn Thr Thr Asp Asn Phe Phe
                   2070                    2075                    2080
    AAA ATT TCC AAC CTG AAG ATG GGT CAC AAT TAC ACA TTC ACG GTC CAG   6468
    Lys Ile Ser Asn Leu Lys Met Gly His Asn Tyr Thr Phe Thr Val Gln
              2085                    2090                    2095
    GCA CGA TGC CTT TTG GGC AGC CAG ATC TGC GGG GAG CCT GCC GTG CTA   6516
    Ala Arg Cys Leu Leu Gly Ser Gln Ile Cys Gly Glu Pro Ala Val Leu
              2100                    2105                    2110
    CTG TAT GAT GAG CTG GGG TCT GGT GGC GAT GCG TCG GCG ATG CAG GCT   6564
    Leu Tyr Asp Glu Leu Gly Ser Gly Gly Asp Ala Ser Ala Met Gln Ala
    2115                    2120                    2125
    GCC AGG TCT ACT GAT GTC GCC GCC GTG GTG GTG CCC ATC CTG TTT CTG   6612
    Ala Arg Ser Thr Asp Val Ala Ala Val Val Val Pro Ile Leu Phe Leu
    2130                    2135                    2140         2145
    ATA CTG CTG AGC CTG GGG GTC GGG TTT GCC ATC CTG TAC ACG AAG CAT   6660
    Ile Leu Leu Ser Leu Gly Val Gly Phe Ala Ile Leu Tyr Thr Lys His
                   2150                    2155                    2160
    CGG AGG CTG CAG AGC AGC TTC ACC GCC TTC GCC AAC AGC CAC TAC AGC   6708
    Arg Arg Leu Gln Ser Ser Phe Thr Ala Phe Ala Asn Ser His Tyr Ser
                   2165                    2170                    2175
    TCC AGA CTC GGC TCC GCC ATC TTC TCC TCT GGG GAT GAC TTG GGG GAG   6756
    Ser Arg Leu Gly Ser Ala Ile Phe Ser Ser Gly Asp Asp Leu Gly Glu
                   2180                    2185                    2190
    GAT GAT GAA GAT GCT CCT ATG ATC ACT GGA TTT TCG GAC GAC GTC CCC   6804
    Asp Asp Glu Asp Ala Pro Met Ile Thr Gly Phe Ser Asp Asp Val Pro
    2195                    2200                    2205
    ATG GTG ATA GCC TGAAAGAGCT TTCCTCACTA GAAACCAAAT GGTGTAAATA      6856
    Met Val Ile Ala
    2210
    TTTTATTTGA TAAAGATAGT TGATGGTTTA TTTTAAAAGA TGCACTTTGA GTTGCAATAT 6916
    GTTATTTTTA TATGGGCCAA AAACAAAAGC AAAAAAAAAA AAAAA               6961
```

(2) INFORMATION FOR SEQ ID NO: 4:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 300 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: double
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: cDNA to mRNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:

```
ATATCCACAT TGACAGCTAT GGTGAAAATT ATCTAAGCTT CACCCTGACC ATGGAGAGTG      60
ATATCAAGGT GAATGGCTAT GTGGTGAACC TTTTCTGGGC ATTTGACACC CACAAGCAAG     120
AGAGGAGAAC TTTGAACTTC CGAGGAAGCA TATTGTCACA CAAAGTTGGC AATCTGACAG     180
CTCATACATC CTATGAGATT TCTGCCTGGG CCAAGACTGA CTTGGGGGAT AGCCCTCTGG     240
CATTTGAGCA TGTTATGACC AGAGGGGTTC GCCCACCTGC ACCTAGCCTC AAGGCCAAAG     300
```

(2) INFORMATION FOR SEQ ID NO: 5:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 6642 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: double
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: cDNA to mRNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 5:

```
ATGGCGACAC GGAGCAGCAG GAGGGAGTCG CGACTCCCGT TCCTATTCAC CCTGGTCGCA      60
CTGCTGCCGC CCGGAGCTCT CTGCGAAGTC TGGACGCAGA GGCTGCACGG CGGCAGCGCG     120
CCCTTGCCCC AGGACCGGGG CTTCCTCGTG GTGCAGGGCG ACCCGCGCGA GCTGCGGCTG     180
TGGGCGCGCG GGGATGCCAG GGGGGCGAGC CGCGCGGACG AGAAGCCGCT CCGGAGGAAA     240
CGGAGCGCTG CCCTGCAGCC CGAGCCCATC AAGGTGTACG GACAGGTTAG TCTGAATGAT     300
TCCCACAATC AGATGGTGGT GCACTGGGCT GGAGAGAAAA GCAACGTGAT CGTGGCCTTG     360
GCCCGAGATA GCCTGGCATT GGCGAGGCCC AAGAGCAGTG ATGTGTACGT GTCTTACGAC     420
TATGGAAAAT CATTCAAGAA AATTTCAGAC AAGTTAAACT TTGGCTTGGG AAATAGGAGT     480
GAAGCTGTTA TCGCCCAGTT CTACCACAGC CCTGCGGACA CAAGCGGTA CATCTTTGCA      540
GACGCTTATG CCCAGTACCT CTGGATCACG TTTGACTTCT GCAACACTCT TCAAGGCTTT     600
TCCATCCCAT TTCGGGCAGC TGATCTCCTC CTACACAGTA AGGCCTCCAA CCTTCTCTTG     660
GGCTTTGACA GGTCCCACCC CAACAAGCAG CTGTGGAAGT CAGATGACTT TGGCCAGACC     720
TGGATCATGA TTCAGGAACA TGTCAAGTCC TTTTCTTGGG AATTGATCC CTATGACAAA     780
CCAAATACCA TCTACATTGA ACGACACGAA CCCTCTGGCT ACTCCACTGT CTTCCGAAGT     840
ACAGATTTCT TCCAGTCCCG GGAAAACCAG GAAGTGATCC TTGAGGAAGT GAGAGATTTT     900
CAGCTTCGGG ACAAGTACAT GTTTGCTACA AAGGTGGTGC ATCTCTTGGG CAGTGAACAG     960
CAGTCTTCTG TCCAGCTCTG GGTCTCCTTT GGCCGGAAGC CCATGAGAGC AGCCCAGTTT    1020
GTCACAAGAC ATCCTATTAA TGAATATTAC ATCGCAGATG CCTCCGAGGA CCAGGTGTTT    1080
```

```
GTGTGTGTCA GCCACAGTAA CAACCGCACC AATTTATACA TCTCAGAGGC AGAGGGGCTG    1140

AAGTTCTCCC TGTCCTTGGA GAACGTGCTC TATTACAGCC CAGGAGGGGC CGGCAGTGAC    1200

ACCTTGGTGA GGTATTTTGC AAATGAACCA TTTGCTGACT TCCACCGAGT GGAAGGATTG    1260

CAAGGAGTCT ACATTGCTAC TCTGATTAAT GGTTCTATGA ATGAGGAGAA CATGAGATCG    1320

GTCATCACCT TTGACAAAGG GGGAACCTGG GAGTTTCTTC AGGCTCCAGC CTTCACGGGA    1380

TATGGAGAGA AAATCAATTG TGAGCTTTCC CAGGGCTGTT CCCTTCATCT GGCTCAGCGC    1440

CTCAGTCAGC TCCTCAACCT CCAGCTCCGG AGAATGCCCA TCCTGTCCAA GGAGTCGGCT    1500

CCAGGCCTCA TCATCGCCAC TGGCTCAGTG GGAAAGAACT TGGCTAGCAA GACAAACGTG    1560

TACATCTCTA GCAGTGCTGG AGCCAGGTGG CGAGAGGCAC TTCCTGGACC TCACTACTAC    1620

ACATGGGGAG ACCACGGCGG AATCATCACG GCCATTGCCC AGGGCATGGA AACCAACGAG    1680

CTAAAATACA GTACCAATGA AGGGGAGACC TGGAAAACAT TCATCTTCTC TGAGAAGCCA    1740

GTGTTTGTGT ATGGCCTCCT CACAGAACCT GGGGAGAAGA GCACTGTCTT CACCATCTTT    1800

GGCTCGAACA AAGAGAATGT CCACAGCTGG CTGATCCTCC AGGTCAATGC CACGGATGCC    1860

TTGGGAGTTC CCTGCACAGA GAATGACTAC AAGCTGTGGT CACCATCTGA TGAGCGGGGG    1920

AATGAGTGTT TGCTGGGACA CAAGACTGTT TTCAAACGGC GGACCCCCCA TGCCACATGC    1980

TTCAATGGAG AGGACTTTGA CAGGCCGGTG GTCGTGTCCA ACTGCTCCTG CACCCGGGAG    2040

GACTATGAGT GTGACTTCGG TTTCAAGATG AGTGAAGATT TGTCATTAGA GGTTTGTGTT    2100

CCAGATCCGG AATTTTCTGG AAAGTCATAC TCCCCTCCTG TGCCTTGCCC TGTGGGTTCT    2160

ACTTACAGGA GAACGAGAGG CTACCGGAAG ATTTCTGGGG ACACTTGTAG CGGAGGAGAT    2220

GTTGAAGCGC GACTGGAAGG AGAGCTGGTC CCCTGTCCCC TGGCAGAAGA GAACGAGTTC    2280

ATTCTGTATG CTGTGAGGAA ATCCATCTAC CGCTATGACC TGGCCTCGGG AGCCACCGAG    2340

CAGTTGCCTC TCACCGGGCT ACGGGCAGCA GTGGCCCTGG ACTTTGACTA TGAGCACAAC    2400

TGTTTGTATT GGTCCGACCT GGCCTTGGAC GTCATCCAGC GCCTCTGTTT GAATGGAAGC    2460

ACAGGGCAAG AGGTGATCAT CAATTCTGGC CTGGAGACAG TAGAAGCTTT GGCTTTTGAA    2520

CCCCTCAGCC AGCTGCTTTA CTGGGTAGAT GCAGGCTTCA AAAAGATTGA GGTAGCTAAT    2580

CCAGATGGCG ACTTCCGACT CACAATCGTC AATTCCTCTG TGCTTGATCG TCCCAGGGCT    2640

CTGGTCCTCG TGCCCCAAGA GGGGGTGATG TTCTGGACAG ACTGGGGAGA CCTGAAGCCT    2700

GGGATTTATC GGAGCAATAT GGATGGTTCT GCTGCCTATC ACCTGGTGTC TGAGGATGTG    2760

AAGTGGCCCA ATGGCATCTC TGTGGACGAC CAGTGGATTT ACTGGACGGA TGCCTACCTG    2820

GAGTGCATAG AGCGGATCAC GTTCAGTGGC CAGCAGCGCT CTGTCATTCT GGACAACCTC    2880

CCGCACCCCT ATGCCATTGC TGTCTTTAAG AATGAAATCT ACTGGGATGA CTGGTCACAG    2940

CTCAGCATAT TCCGAGCTTC CAAATACAGT GGGTCCCAGA TGGAGATTCT GGCAAACCAG    3000
```

```
CTCACGGGGC TCATGGACAT GAAGATTTTC TACAAGGGGA AGAACACTGG AAGCAATGCC    3060
TGTGTGCCCA GGCCATGCAG CCTGCTGTGC CTGCCCAAGG CCAACAACAG TAGAAGCTGC    3120
AGGTGTCCAG AGGATGTGTC CAGCAGTGTG CTTCCATCAG GGGACCTGAT GTGTGACTGC    3180
CCTCAGGGCT ATCAGCTCAA GAACAATACC TGTGTCAAAG AAGAGAACAC CTGTCTTCGC    3240
AACCAGTATC GCTGCAGCAA CGGGAACTGT ATCAACAGCA TTTGGTGGTG TGACTTTGAC    3300
AACGACTGTG AGACATGAG CGATGAGAGA AACTGCCCTA CCACCATCTG TGACCTGGAC    3360
ACCCAGTTTC GTTGCCAGGA GTCTGGGACT TGTATCCCAC TGTCCTATAA ATGTGACCTT    3420
GAGGATGACT GTGGAGACAA CAGTGATGAA AGTCATTGTG AAATGCACCA GTGCCGGAGT    3480
GACGAGTACA ACTGCAGTTC CGGCATGTGC ATCCGCTCCT CCTGGGTATG TGACGGGGAC    3540
AACGACTGCA GGGACTGGTC TGATGAAGCC AACTGTACCG CCATCTATCA CACCTGTGAG    3600
GCCTCCAACT TCCAGTGCCG AAACGGGCAC TGCATCCCCC AGCGGTGGGC GTGTGACGGG    3660
GATACGGACT GCCAGGATGG TTCCGATGAG GATCCAGTCA ACTGTGAGAA GAAGTGCAAT    3720
GGATTCCGCT GCCCAAACGG CACTTGCATC CCATCCAGCA AACATTGTGA TGGTCTGCGT    3780
GATTGCTCTG ATGGCTCCGA TGAACAGCAC TGCGAGCCCC TCTGTACGCA CTTCATGGAC    3840
TTTGTGTGTA AGAACCGCCA GCAGTGCCTG TTCCACTCCA TGGTCTGTGA CGGAATCATC    3900
CAGTGCCGCG ACGGGTCCGA TGAGGATGCG GCGTTTGCAG GATGCTCCCA AGATCCTGAG    3960
TTCCACAAGG TATGTGATGA GTTCGGTTTC CAGTGTCAGA ATGGAGTGTG CATCAGTTTG    4020
ATTTGGAAGT GCGACGGGAT GGATGATTGC GGCGATTATT CTGATGAAGC CAACTGCGAA    4080
AACCCCACAG AAGCCCCAAA CTGCTCCCGC TACTTCCAGT TTCGGTGTGA GAATGGCCAC    4140
TGCATCCCCA ACAGATGGAA ATGTGACAGG GAGAACGACT GTGGGGACTG GTCTGATGAG    4200
AAGGATTGTG GAGATTCACA TATTCTTCCC TTCTCGACTC CTGGGCCCTC CACGTGTCTG    4260
CCCAATTACT ACCGCTGCAG CAGTGGGACC TGCGTGATGG ACACCTGGGT GTGCGACGGG    4320
TACCGAGATT GTGCAGATGG CTCTGACGAG GAAGCCTGCC CCTTGCTTGC AAACGTCACT    4380
GCTGCCTCCA CTCCCACCCA ACTTGGGCGA TGTGACCGAT TTGAGTTCGA ATGCCACCAA    4440
CCGAAGACGT GTATTCCCAA CTGGAAGCGC TGTGACGGCC ACCAAGATTG CCAGGATGGC    4500
CGGGACGAGG CCAATTGCCC CACACACAGC ACCTTGACTT GCATGAGCAG GGAGTTCCAG    4560
TGCGAGGACG GGGAGGCCTG CATTGTGCTC TCGGAGCGCT GCGACGGCTT CCTGGACTGC    4620
TCGGACGAGA GCGATGAAAA GGCCTGCAGT GATGAGTTGA CTGTGTACAA AGTACAGAAT    4680
CTTCAGTGGA CAGCTGACTT CTCTGGGGAT GTGACTTTGA CCTGGATGAG GCCCAAAAAA    4740
ATGCCCTCTG CATCTTGTGT ATATAATGTC TACTACAGGG TGGTTGGAGA GAGCATATGG    4800
AAGACTCTGG AGACCCACAG CAATAAGACA AACACTGTAT TAAAAGTCTT GAAACCAGAT    4860
ACCACGTATC AGGTTAAAGT ACAGGTTCAG TGTCTCAGCA AGGCACACAA CACCAATGAC    4920
```

```
TTTGTGACCC TGAGGACCCC AGAGGGATTG CCAGATGCCC CTCGAAATCT CCAGCTGTCA    4980

CTCCCCAGGG AAGCAGAAGG TGTGATTGTA GGCCACTGGG CTCCTCCCAT CCACACCCAT    5040

GGCCTCATCC GTGAGTACAT TGTAGAATAC AGCAGGAGTG GTTCCAAGAT GTGGGCCTCC    5100

CAGAGGGCTG CTAGTAACTT TACAGAAATC AAGAACTTAT TGGTCAACAC TCTATACACC    5160

GTCAGAGTGG CTGCGGTGAC TAGTCGTGGA ATAGGAAACT GGAGCGATTC TAAATCCATT    5220

ACCACCATAA AAGGAAAAGT GATCCCACCA CCAGATATCC ACATTGACAG CTATGGTGAA    5280

AATTATCTAA GCTTCACCCT GACCATGGAG AGTGATATCA AGGTGAATGG CTATGTGGTG    5340

AACCTTTTCT GGGCATTTGA CACCCACAAG CAAGAGAGGA GAACTTTGAA CTTCCGAGGA    5400

AGCATATTGT CACACAAAGT TGGCAATCTG ACAGCTCATA CATCCTATGA GATTTCTGCC    5460

TGGGCCAAGA CTGACTTGGG GGATAGCCCT CTGGCATTTG AGCATGTTAT GACCAGAGGG    5520

GTTCGCCCAC CTGCACCTAG CCTCAAGGCC AAAGCCATCA ACCAGACTGC AGTGGAATGT    5580

ACCTGGACCG GCCCCCGGAA TGTGGTTTAT GGTATTTTCT ATGCCACGTC CTTTCTTGAC    5640

CTCTATCGCA ACCCGAAGAG CTTGACTACT TCACTCCACA ACAAGACGGT CATTGTCAGT    5700

AAGGATGAGC AGTATTTGTT TCTGGTCCGT GTAGTGGTAC CCTACCAGGG GCCATCCTCT    5760

GACTACGTTG TAGTGAAGAT GATCCCGGAC AGCAGGCTTC CACCCCGTCA CCTGCATGTG    5820

GTTCATACGG GCAAAACCTC CGTGGTCATC AAGTGGGAAT CACCGTATGA CTCTCCTGAC    5880

CAGGACTTGT TGTATGCAAT TGCAGTCAAA GATCTCATAA GAAAGACTGA CAGGAGCTAC    5940

AAAGTAAAAT CCCGTAACAG CACTGTGGAA TACACCCTTA ACAAGTTGGA GCCTGGCGGG    6000

AAATACCACA TCATTGTCCA ACTGGGGAAC ATGAGCAAAG ATTCCAGCAT AAAAATTACC    6060

ACAGTTTCAT TATCAGCACC TGATGCCTTA AAAATCATAA CAGAAAATGA TCATGTTCTT    6120

CTGTTTTGGA AAAGCCTGGC TTTAAAGGAA AAGCATTTTA ATGAAAGCAG GGGCTATGAG    6180

ATACACATGT TTGATAGTGC CATGAATATC ACAGCTTACC TTGGGAATAC TACTGACAAT    6240

TTCTTTAAAA TTTCCAACCT GAAGATGGGT CATAATTACA CGTTCACCGT CCAAGCAAGA    6300

TGCCTTTTTG GCAACCAGAT CTGTGGGGAG CCTGCCATCC TGCTGTACGA TGAGCTGGGG    6360

TCTGGTGCAG ATGCATCTGC AACGCAGGCT GCCAGATCTA CGGATGTTGC TGCTGTGGTG    6420

GTGCCCATCT TATTCCTGAT ACTGCTGAGC CTGGGGGTGG GGTTTGCCAT CCTGTACACG    6480

AAGCACCGGA GGCTGCAGAG CAGCTTCACC GCCTTCGCCA ACAGCCACTA CAGCTCCAGG    6540

CTGGGGTCCG CAATCTTCTC CTCTGGGGAT GACCTGGGGG AAGATGATGA AGATGCCCCT    6600

ATGATAACTG GATTTTCAGA TGACGTCCCC ATGGTGATAG CC                       6642
```

(2) INFORMATION FOR SEQ ID NO: 6:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 2214 amino acids
        (B) TYPE: amino acid

(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 6:

```
Met Ala Thr Arg Ser Ser Arg Arg Glu Ser Arg Leu Pro Phe Leu Phe
1               5               10              15

Thr Leu Val Ala Leu Leu Pro Pro Gly Ala Leu Cys Glu Val Trp Thr
            20              25              30

Gln Arg Leu His Gly Gly Ser Ala Pro Leu Pro Gln Asp Arg Gly Phe
        35              40              45

Leu Val Val Gln Gly Asp Pro Arg Glu Leu Arg Leu Trp Ala Arg Gly
        50              55              60

Asp Ala Arg Gly Ala Ser Arg Ala Asp Glu Lys Pro Leu Arg Arg Lys
65              70              75              80

Arg Ser Ala Ala Leu Gln Pro Glu Pro Ile Lys Val Tyr Gly Gln Val
                85              90              95

Ser Leu Asn Asp Ser His Asn Gln Met Val Val His Trp Ala Gly Glu
            100             105             110

Lys Ser Asn Val Ile Val Ala Leu Ala Arg Asp Ser Leu Ala Leu Ala
            115             120             125

Arg Pro Lys Ser Ser Asp Val Tyr Val Ser Tyr Asp Tyr Gly Lys Ser
    130             135             140

Phe Lys Lys Ile Ser Asp Lys Leu Asn Phe Gly Leu Gly Asn Arg Ser
145             150             155             160

Glu Ala Val Ile Ala Gln Phe Tyr His Ser Pro Ala Asp Asn Lys Arg
            165             170             175

Tyr Ile Phe Ala Asp Ala Tyr Ala Gln Tyr Leu Trp Ile Thr Phe Asp
            180             185             190

Phe Cys Asn Thr Leu Gln Gly Phe Ser Ile Pro Phe Arg Ala Ala Asp
        195             200             205

Leu Leu Leu His Ser Lys Ala Ser Asn Leu Leu Leu Gly Phe Asp Arg
    210             215             220

Ser His Pro Asn Lys Gln Leu Trp Lys Ser Asp Asp Phe Gly Gln Thr
225             230             235             240

Trp Ile Met Ile Gln Glu His Val Lys Ser Phe Ser Trp Gly Ile Asp
            245             250             255

Pro Tyr Asp Lys Pro Asn Thr Ile Tyr Ile Glu Arg His Glu Pro Ser
            260             265             270

Gly Tyr Ser Thr Val Phe Arg Ser Thr Asp Phe Phe Gln Ser Arg Glu
        275             280             285
```

55

90

```
Asn Gln Glu Val Ile Leu Glu Glu Val Arg Asp Phe Gln Leu Arg Asp
    290                 295             300
```

5
```
Lys Tyr Met Phe Ala Thr Lys Val Val His Leu Leu Gly Ser Glu Gln
305             310             315             320
```

```
Gln Ser Ser Val Gln Leu Trp Val Ser Phe Gly Arg Lys Pro Met Arg
            325             330             335
```

10
```
Ala Ala Gln Phe Val Thr Arg His Pro Ile Asn Glu Tyr Tyr Ile Ala
        340             345             350
```

```
Asp Ala Ser Glu Asp Gln Val Phe Val Cys Val Ser His Ser Asn Asn
        355             360             365
```

15
```
Arg Thr Asn Leu Tyr Ile Ser Glu Ala Glu Gly Leu Lys Phe Ser Leu
    370             375             380
```

```
Ser Leu Glu Asn Val Leu Tyr Tyr Ser Pro Gly Gly Ala Gly Ser Asp
385             390             395             400
```

20
```
Thr Leu Val Arg Tyr Phe Ala Asn Glu Pro Phe Ala Asp Phe His Arg
            405             410             415
```

```
Val Glu Gly Leu Gln Gly Val Tyr Ile Ala Thr Leu Ile Asn Gly Ser
            420             425             430
```

25
```
Met Asn Glu Glu Asn Met Arg Ser Val Ile Thr Phe Asp Lys Gly Gly
        435             440             445
```

```
Thr Trp Glu Phe Leu Gln Ala Pro Ala Phe Thr Gly Tyr Gly Glu Lys
    450             455             460
```

30
```
Ile Asn Cys Glu Leu Ser Gln Gly Cys Ser Leu His Leu Ala Gln Arg
465             470             475             480
```

```
Leu Ser Gln Leu Leu Asn Leu Gln Leu Arg Arg Met Pro Ile Leu Ser
            485             490             495
```

35
```
Lys Glu Ser Ala Pro Gly Leu Ile Ile Ala Thr Gly Ser Val Gly Lys
            500             505             510
```

```
Asn Leu Ala Ser Lys Thr Asn Val Tyr Ile Ser Ser Ser Ala Gly Ala
        515             520             525
```

40
```
Arg Trp Arg Glu Ala Leu Pro Gly Pro His Tyr Tyr Thr Trp Gly Asp
    530             535             540
```

```
His Gly Gly Ile Ile Thr Ala Ile Ala Gln Gly Met Glu Thr Asn Glu
545             550             555             560
```

45
```
Leu Lys Tyr Ser Thr Asn Glu Gly Glu Thr Trp Lys Thr Phe Ile Phe
            565             570             575
```

```
Ser Glu Lys Pro Val Phe Val Tyr Gly Leu Leu Thr Glu Pro Gly Glu
            580             585             590
```

```
Lys Ser Thr Val Phe Thr Ile Phe Gly Ser Asn Lys Glu Asn Val His
            595             600             605
```

50
```
Ser Trp Leu Ile Leu Gln Val Asn Ala Thr Asp Ala Leu Gly Val Pro
    610             615             620
```

```
Cys Thr Glu Asn Asp Tyr Lys Leu Trp Ser Pro Ser Asp Glu Arg Gly
```

55

```
        625                    630                    635                    640

        Asn Glu Cys Leu Leu Gly His Lys Thr Val Phe Lys Arg Arg Thr Pro
5                       645             650                 655

        His Ala Thr Cys Phe Asn Gly Glu Asp Phe Asp Arg Pro Val Val Val
                    660             665                 670

        Ser Asn Cys Ser Cys Thr Arg Glu Asp Tyr Glu Cys Asp Phe Gly Phe
10                  675             680                 685

        Lys Met Ser Glu Asp Leu Ser Leu Glu Val Cys Val Pro Asp Pro Glu
                690             695                 700

        Phe Ser Gly Lys Ser Tyr Ser Pro Pro Val Pro Cys Pro Val Gly Ser
        705                 710                 715                 720

        Thr Tyr Arg Arg Thr Arg Gly Tyr Arg Lys Ile Ser Gly Asp Thr Cys
15                      725             730                 735

        Ser Gly Gly Asp Val Glu Ala Arg Leu Glu Gly Glu Leu Val Pro Cys
                    740             745                 750

        Pro Leu Ala Glu Glu Asn Glu Phe Ile Leu Tyr Ala Val Arg Lys Ser
20                  755             760                 765

        Ile Tyr Arg Tyr Asp Leu Ala Ser Gly Ala Thr Glu Gln Leu Pro Leu
                770             775                 780

        Thr Gly Leu Arg Ala Ala Val Ala Leu Asp Phe Asp Tyr Glu His Asn
25      785             790             795                 800

        Cys Leu Tyr Trp Ser Asp Leu Ala Leu Asp Val Ile Gln Arg Leu Cys
                    805             810                 815

        Leu Asn Gly Ser Thr Gly Gln Glu Val Ile Ile Asn Ser Gly Leu Glu
30                  820             825                 830

        Thr Val Glu Ala Leu Ala Phe Glu Pro Leu Ser Gln Leu Leu Tyr Trp
                    835             840                 845

        Val Asp Ala Gly Phe Lys Lys Ile Glu Val Ala Asn Pro Asp Gly Asp
35          850             855                 860

        Phe Arg Leu Thr Ile Val Asn Ser Ser Val Leu Asp Arg Pro Arg Ala
        865                 870             875                 880

        Leu Val Leu Val Pro Gln Glu Gly Val Met Phe Trp Thr Asp Trp Gly
40                  885             890                 895

        Asp Leu Lys Pro Gly Ile Tyr Arg Ser Asn Met Asp Gly Ser Ala Ala
                    900             905                 910

        Tyr His Leu Val Ser Glu Asp Val Lys Trp Pro Asn Gly Ile Ser Val
45          915             920                 925

        Asp Asp Gln Trp Ile Tyr Trp Thr Asp Ala Tyr Leu Glu Cys Ile Glu
                930             935                 940

        Arg Ile Thr Phe Ser Gly Gln Gln Arg Ser Val Ile Leu Asp Asn Leu
50      945             950                 955                 960

        Pro His Pro Tyr Ala Ile Ala Val Phe Lys Asn Glu Ile Tyr Trp Asp
                    965             970                 975

55
```

```
Asp Trp Ser Gln Leu Ser Ile Phe Arg Ala Ser Lys Tyr Ser Gly Ser
         980                 985                 990

Gln Met Glu Ile Leu Ala Asn Gln Leu Thr Gly Leu Met Asp Met Lys
         995                1000                1005

Ile Phe Tyr Lys Gly Lys Asn Thr Gly Ser Asn Ala Cys Val Pro Arg
        1010                1015                1020

Pro Cys Ser Leu Leu Cys Leu Pro Lys Ala Asn Asn Ser Arg Ser Cys
1025                1030                1035                1040

Arg Cys Pro Glu Asp Val Ser Ser Ser Val Leu Pro Ser Gly Asp Leu
                1045                1050                1055

Met Cys Asp Cys Pro Gln Gly Tyr Gln Leu Lys Asn Asn Thr Cys Val
                1060                1065                1070

Lys Glu Glu Asn Thr Cys Leu Arg Asn Gln Tyr Arg Cys Ser Asn Gly
        1075                1080                1085

Asn Cys Ile Asn Ser Ile Trp Trp Cys Asp Phe Asp Asn Asp Cys Gly
        1090                1095                1100

Asp Met Ser Asp Glu Arg Asn Cys Pro Thr Thr Ile Cys Asp Leu Asp
1105                1110                1115                1120

Thr Gln Phe Arg Cys Gln Glu Ser Gly Thr Cys Ile Pro Leu Ser Tyr
                1125                1130                1135

Lys Cys Asp Leu Glu Asp Asp Cys Gly Asp Asn Ser Asp Glu Ser His
                1140                1145                1150

Cys Glu Met His Gln Cys Arg Ser Asp Glu Tyr Asn Cys Ser Ser Gly
        1155                1160                1165

Met Cys Ile Arg Ser Ser Trp Val Cys Asp Gly Asp Asn Asp Cys Arg
        1170                1175                1180

Asp Trp Ser Asp Glu Ala Asn Cys Thr Ala Ile Tyr His Thr Cys Glu
1185                1190                1195                1200

Ala Ser Asn Phe Gln Cys Arg Asn Gly His Cys Ile Pro Gln Arg Trp
                1205                1210                1215

Ala Cys Asp Gly Asp Thr Asp Cys Gln Asp Gly Ser Asp Glu Asp Pro
                1220                1225                1230

Val Asn Cys Glu Lys Lys Cys Asn Gly Phe Arg Cys Pro Asn Gly Thr
                1235                1240                1245

Cys Ile Pro Ser Ser Lys His Cys Asp Gly Leu Arg Asp Cys Ser Asp
        1250                1255                1260

Gly Ser Asp Glu Gln His Cys Glu Pro Leu Cys Thr His Phe Met Asp
1265                1270                1275                1280

Phe Val Cys Lys Asn Arg Gln Gln Cys Leu Phe His Ser Met Val Cys
                1285                1290                1295

Asp Gly Ile Ile Gln Cys Arg Asp Gly Ser Asp Glu Asp Ala Ala Phe
                1300                1305                1310
```

Ala Gly Cys Ser Gln Asp Pro Glu Phe His Lys Val Cys Asp Glu Phe
     1315   1320   1325

Gly Phe Gln Cys Gln Asn Gly Val Cys Ile Ser Leu Ile Trp Lys Cys
     1330   1335   1340

Asp Gly Met Asp Asp Cys Gly Asp Tyr Ser Asp Glu Ala Asn Cys Glu
1345   1350   1355   1360

Asn Pro Thr Glu Ala Pro Asn Cys Ser Arg Tyr Phe Gln Phe Arg Cys
          1365   1370   1375

Glu Asn Gly His Cys Ile Pro Asn Arg Trp Lys Cys Asp Arg Glu Asn
         1380   1385   1390

Asp Cys Gly Asp Trp Ser Asp Glu Lys Asp Cys Gly Asp Ser His Ile
     1395   1400   1405

Leu Pro Phe Ser Thr Pro Gly Pro Ser Thr Cys Leu Pro Asn Tyr Tyr
1410   1415   1420

Arg Cys Ser Ser Gly Thr Cys Val Met Asp Thr Trp Val Cys Asp Gly
1425   1430   1435   1440

Tyr Arg Asp Cys Ala Asp Gly Ser Asp Glu Glu Ala Cys Pro Leu Leu
         1445   1450   1455

Ala Asn Val Thr Ala Ala Ser Thr Pro Thr Gln Leu Gly Arg Cys Asp
         1460   1465   1470

Arg Phe Glu Phe Glu Cys His Gln Pro Lys Thr Cys Ile Pro Asn Trp
     1475   1480   1485

Lys Arg Cys Asp Gly His Gln Asp Cys Gln Asp Gly Arg Asp Glu Ala
     1490   1495   1500

Asn Cys Pro Thr His Ser Thr Leu Thr Cys Met Ser Arg Glu Phe Gln
1505   1510   1515   1520

Cys Glu Asp Gly Glu Ala Cys Ile Val Leu Ser Glu Arg Cys Asp Gly
         1525   1530   1535

Phe Leu Asp Cys Ser Asp Glu Ser Asp Glu Lys Ala Cys Ser Asp Glu
     1540   1545   1550

Leu Thr Val Tyr Lys Val Gln Asn Leu Gln Trp Thr Ala Asp Phe Ser
     1555   1560   1565

Gly Asp Val Thr Leu Thr Trp Met Arg Pro Lys Lys Met Pro Ser Ala
     1570   1575   1580

Ser Cys Val Tyr Asn Val Tyr Tyr Arg Val Val Gly Glu Ser Ile Trp
1585   1590   1595   1600

Lys Thr Leu Glu Thr His Ser Asn Lys Thr Asn Thr Val Leu Lys Val
         1605   1610   1615

Leu Lys Pro Asp Thr Thr Tyr Gln Val Lys Val Gln Val Gln Cys Leu
     1620   1625   1630

Ser Lys Ala His Asn Thr Asn Asp Phe Val Thr Leu Arg Thr Pro Glu
     1635   1640   1645

Gly Leu Pro Asp Ala Pro Arg Asn Leu Gln Leu Ser Leu Pro Arg Glu

```
          1650                    1655                   1660

          Ala Glu Gly Val Ile Val Gly His Trp Ala Pro Pro Ile His Thr His
  5       1665                    1670                   1675            1680

          Gly Leu Ile Arg Glu Tyr Ile Val Glu Tyr Ser Arg Ser Gly Ser Lys
                             1685                    1690            1695

          Met Trp Ala Ser Gln Arg Ala Ala Ser Asn Phe Thr Glu Ile Lys Asn
  10                        1700                    1705            1710

          Leu Leu Val Asn Thr Leu Tyr Thr Val Arg Val Ala Ala Val Thr Ser
                        1715                    1720            1725

          Arg Gly Ile Gly Asn Trp Ser Asp Ser Lys Ser Ile Thr Thr Ile Lys
                        1730                    1735            1740

  15      Gly Lys Val Ile Pro Pro Pro Asp Ile His Ile Asp Ser Tyr Gly Glu
          1745                    1750                    1755           1760

          Asn Tyr Leu Ser Phe Thr Leu Thr Met Glu Ser Asp Ile Lys Val Asn
                             1765                    1770            1775

  20      Gly Tyr Val Val Asn Leu Phe Trp Ala Phe Asp Thr His Lys Gln Glu
                        1780                    1785            1790

          Arg Arg Thr Leu Asn Phe Arg Gly Ser Ile Leu Ser His Lys Val Gly
                        1795                    1800            1805

  25      Asn Leu Thr Ala His Thr Ser Tyr Glu Ile Ser Ala Trp Ala Lys Thr
                        1810                    1815            1820

          Asp Leu Gly Asp Ser Pro Leu Ala Phe Glu His Val Met Thr Arg Gly
          1825                    1830                    1835           1840

  30      Val Arg Pro Pro Ala Pro Ser Leu Lys Ala Lys Ala Ile Asn Gln Thr
                             1845                    1850            1855

          Ala Val Glu Cys Thr Trp Thr Gly Pro Arg Asn Val Val Tyr Gly Ile
                        1860                    1865                 1870

  35      Phe Tyr Ala Thr Ser Phe Leu Asp Leu Tyr Arg Asn Pro Lys Ser Leu
                        1875                    1880                 1885

          Thr Thr Ser Leu His Asn Lys Thr Val Ile Val Ser Lys Asp Glu Gln
                        1890                    1895            1900

  40      Tyr Leu Phe Leu Val Arg Val Val Val Pro Tyr Gln Gly Pro Ser Ser
          1905                    1910                    1915           1920

          Asp Tyr Val Val Val Lys Met Ile Pro Asp Ser Arg Leu Pro Pro Arg
                             1925                    1930            1935

  45      His Leu His Val Val His Thr Gly Lys Thr Ser Val Val Ile Lys Trp
                        1940                    1945                 1950

          Glu Ser Pro Tyr Asp Ser Pro Asp Gln Asp Leu Leu Tyr Ala Ile Ala
                        1955                    1960                 1965

  50      Val Lys Asp Leu Ile Arg Lys Thr Asp Arg Ser Tyr Lys Val Lys Ser
                        1970                    1975                 1980

          Arg Asn Ser Thr Val Glu Tyr Thr Leu Asn Lys Leu Glu Pro Gly Gly
          1985                    1990                    1995           2000

  55
```

```
        Lys Tyr His Ile Ile Val Gln Leu Gly Asn Met Ser Lys Asp Ser Ser
                        2005                2010              2015
5       Ile Lys Ile Thr Thr Val Ser Leu Ser Ala Pro Asp Ala Leu Lys Ile
                        2020                2025              2030
        Ile Thr Glu Asn Asp His Val Leu Leu Phe Trp Lys Ser Leu Ala Leu
                        2035                2040              2045
10      Lys Glu Lys His Phe Asn Glu Ser Arg Gly Tyr Glu Ile His Met Phe
                        2050                2055              2060
        Asp Ser Ala Met Asn Ile Thr Ala Tyr Leu Gly Asn Thr Thr Asp Asn
        2065                2070                2075              2080
15      Phe Phe Lys Ile Ser Asn Leu Lys Met Gly His Asn Tyr Thr Phe Thr
                        2085                2090              2095
        Val Gln Ala Arg Cys Leu Phe Gly Asn Gln Ile Cys Gly Glu Pro Ala
                        2100                2105              2110
20      Ile Leu Leu Tyr Asp Glu Leu Gly Ser Gly Ala Asp Ala Ser Ala Thr
                        2115                2120              2125
        Gln Ala Ala Arg Ser Thr Asp Val Ala Ala Val Val Val Pro Ile Leu
                        2130                2135              2140
25      Phe Leu Ile Leu Leu Ser Leu Gly Val Gly Phe Ala Ile Leu Tyr Thr
        2145                2150                2155              2160
        Lys His Arg Arg Leu Gln Ser Ser Phe Thr Ala Phe Ala Asn Ser His
                        2165                2170              2175
30      Tyr Ser Ser Arg Leu Gly Ser Ala Ile Phe Ser Ser Gly Asp Asp Leu
                        2180                2185              2190
        Gly Glu Asp Asp Glu Asp Ala Pro Met Ile Thr Gly Phe Ser Asp Asp
                        2195                2200              2205
35      Val Pro Met Val Ile Ala
                        2210
```

(2) INFORMATION FOR SEQ ID NO: 7:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 6843 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: double
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: cDNA to mRNA

    (ix) FEATURE:
        (A) NAME/KEY: sig peptide
        (B) LOCATION:81..164
        (C) IDENTIFICATION METHOD: S

    (ix) FEATURE:
        (A) NAME/KEY: mat peptide
        (B) LOCATION:165..6722
        (C) IDENTIFICATION METHOD: S

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 7:

```
                                        CCG GCCCAGCGGC TCTCCTGGCC    23
TCGCGCTGCA CATTCTCTCC TGGCGGCGGC GCCACCTGCA GTAGCGTTCG CCCGAACATG    83
                                                              Met
                                                               1
GCG ACA CGG AGC AGC AGG AGG GAG TCG CGA CTC CCG TTC CTA TTC ACC   131
Ala Thr Arg Ser Ser Arg Arg Glu Ser Arg Leu Pro Phe Leu Phe Thr
              5                   10                  15
CTG GTC GCA CTG CTG CCG CCC GGA GCT CTC TGC GAA GTC TGG ACG CAG   179
Leu Val Ala Leu Leu Pro Pro Gly Ala Leu Cys Glu Val Trp Thr Gln
          20                  25                  30
AGG CTG CAC GGC GGC AGC GCG CCC TTG CCC CAG GAC CGG GGC TTC CTC   227
Arg Leu His Gly Gly Ser Ala Pro Leu Pro Gln Asp Arg Gly Phe Leu
      35                  40                  45
GTG GTG CAG GGC GAC CCG CGC GAG CTG CGG CTG TGG GCG CGC GGG GAT   275
Val Val Gln Gly Asp Pro Arg Glu Leu Arg Leu Trp Ala Arg Gly Asp
  50                  55                  60                  65
GCC AGG GGG GCG AGC CGC GCG GAC GAG AAG CCG CTC CGG AGG AAA CGG   323
Ala Arg Gly Ala Ser Arg Ala Asp Glu Lys Pro Leu Arg Arg Lys Arg
                  70                  75                  80
AGC GCT GCC CTG CAG CCC GAG CCC ATC AAG GTG TAC GGA CAG GTT AGT   371
Ser Ala Ala Leu Gln Pro Glu Pro Ile Lys Val Tyr Gly Gln Val Ser
              85                  90                  95
CTG AAT GAT TCC CAC AAT CAG ATG GTG GTG CAC TGG GCT GGA GAG AAA   419
Leu Asn Asp Ser His Asn Gln Met Val Val His Trp Ala Gly Glu Lys
          100                 105                 110
AGC AAC GTG ATC GTG GCC TTG GCC CGA GAT AGC CTG GCA TTG GCG AGG   467
Ser Asn Val Ile Val Ala Leu Ala Arg Asp Ser Leu Ala Leu Ala Arg
          115                 120                 125
CCC AAG AGC AGT GAT GTG TAC GTG TCT TAC GAC TAT GGA AAA TCA TTC   515
Pro Lys Ser Ser Asp Val Tyr Val Ser Tyr Asp Tyr Gly Lys Ser Phe
130                 135                 140                 145
AAG AAA ATT TCA GAC AAG TTA AAC TTT GGC TTG GGA AAT AGG AGT GAA   563
Lys Lys Ile Ser Asp Lys Leu Asn Phe Gly Leu Gly Asn Arg Ser Glu
                  150                 155                 160
GCT GTT ATC GCC CAG TTC TAC CAC AGC CCT GCG GAC AAC AAG CGG TAC   611
Ala Val Ile Ala Gln Phe Tyr His Ser Pro Ala Asp Asn Lys Arg Tyr
              165                 170                 175
ATC TTT GCA GAC GCT TAT GCC CAG TAC CTC TGG ATC ACG TTT GAC TTC   659
Ile Phe Ala Asp Ala Tyr Ala Gln Tyr Leu Trp Ile Thr Phe Asp Phe
              180                 185                 190
TGC AAC ACT CTT CAA GGC TTT TCC ATC CCA TTT CGG GCA GCT GAT CTC   707
Cys Asn Thr Leu Gln Gly Phe Ser Ile Pro Phe Arg Ala Ala Asp Leu
          195                 200                 205
CTC CTA CAC AGT AAG GCC TCC AAC CTT CTC TTG GGC TTT GAC AGG TCC   755
Leu Leu His Ser Lys Ala Ser Asn Leu Leu Leu Gly Phe Asp Arg Ser
210                 215                 220                 225
CAC CCC AAC AAG CAG CTG TGG AAG TCA GAT GAC TTT GGC CAG ACC TGG   803
His Pro Asn Lys Gln Leu Trp Lys Ser Asp Asp Phe Gly Gln Thr Trp
                  230                 235                 240
ATC ATG ATT CAG GAA CAT GTC AAG TCC TTT TCT TGG GGA ATT GAT CCC   851
Ile Met Ile Gln Glu His Val Lys Ser Phe Ser Trp Gly Ile Asp Pro
                  245                 250                 255
TAT GAC AAA CCA AAT ACC ATC TAC ATT GAA CGA CAC GAA CCC TCT GGC   899
Tyr Asp Lys Pro Asn Thr Ile Tyr Ile Glu Arg His Glu Pro Ser Gly
              260                 265                 270
TAC TCC ACT GTC TTC CGA AGT ACA GAT TTC TTC CAG TCC CGG GAA AAC   947
Tyr Ser Thr Val Phe Arg Ser Thr Asp Phe Phe Gln Ser Arg Glu Asn
          275                 280                 285
CAG GAA GTG ATC CTT GAG GAA GTG AGA GAT TTT CAG CTT CGG GAC AAG   995
Gln Glu Val Ile Leu Glu Glu Val Arg Asp Phe Gln Leu Arg Asp Lys
290                 295                 300                 305
```

```
TAC ATG TTT GCT ACA AAG GTG GTG CAT CTC TTG GGC ACT GAA CAC CAG    1043
Tyr Met Phe Ala Thr Lys Val Val His Leu Leu Gly Ser Glu Gln Gln
            310                     315                 320
TCT TCT GTC CAG CTC TGG GTC TCC TTT GGC CGG AAG CCC ATG AGA GCA    1091
Ser Ser Val Gln Leu Trp Val Ser Phe Gly Arg Lys Pro Met Arg Ala
            325                     330                 335
GCC CAG TTT GTC ACA AGA CAT CCT ATT AAT GAA TAT TAC ATC GCA GAT    1139
Ala Gln Phe Val Thr Arg His Pro Ile Asn Glu Tyr Tyr Ile Ala Asp
            340                     345                 350
GCC TCC GAG GAC CAG GTG TTT GTG TGT GTC AGC CAC AGT AAC AAC CGC    1187
Ala Ser Glu Asp Gln Val Phe Val Cys Val Ser His Ser Asn Asn Arg
        355                     360                 365
ACC AAT TTA TAC ATC TCA GAG GCA GAG GGG CTG AAG TTC TCC CTG TCC    1235
Thr Asn Leu Tyr Ile Ser Glu Ala Glu Gly Leu Lys Phe Ser Leu Ser
370                     375                     380             385
TTG GAG AAC GTG CTC TAT TAC AGC CCA GGA GGG GCC GGC AGT GAC ACC    1283
Leu Glu Asn Val Leu Tyr Tyr Ser Pro Gly Gly Ala Gly Ser Asp Thr
                390                     395                 400
TTG GTG AGG TAT TTT GCA AAT GAA CCA TTT GCT GAC TTC CAC CGA GTG    1331
Leu Val Arg Tyr Phe Ala Asn Glu Pro Phe Ala Asp Phe His Arg Val
                405                     410                 415
GAA GGA TTG CAA GGA GTC TAC ATT GCT ACT CTG ATT AAT GGT TCT ATG    1379
Glu Gly Leu Gln Gly Val Tyr Ile Ala Thr Leu Ile Asn Gly Ser Met
            420                     425                 430
AAT GAG GAG AAC ATG AGA TCG GTC ATC ACC TTT GAC AAA GGG GGA ACC    1427
Asn Glu Glu Asn Met Arg Ser Val Ile Thr Phe Asp Lys Gly Gly Thr
        435                     440                 445
TGG GAG TTT CTT CAG GCT CCA GCC TTC ACG GGA TAT GGA GAG AAA ATC    1475
Trp Glu Phe Leu Gln Ala Pro Ala Phe Thr Gly Tyr Gly Glu Lys Ile
450                     455                 460                 465
AAT TGT GAG CTT TCC CAG GGC TGT TCC CTT CAT CTG GCT CAG CGC CTC    1523
Asn Cys Glu Leu Ser Gln Gly Cys Ser Leu His Leu Ala Gln Arg Leu
                470                     475                 480
AGT CAG CTC CTC AAC CTC CAG CTC CGG AGA ATG CCC ATC CTG TCC AAG    1571
Ser Gln Leu Leu Asn Leu Gln Leu Arg Arg Met Pro Ile Leu Ser Lys
            485                     490                 495
GAG TCG GCT CCA GGC CTC ATC ATC GCC ACT GGC TCA GTG GGA AAG AAC    1619
Glu Ser Ala Pro Gly Leu Ile Ile Ala Thr Gly Ser Val Gly Lys Asn
            500                     505                 510
TTG GCT AGC AAG ACA AAC GTG TAC ATC TCT AGC AGT GCT GGA GCC AGG    1667
Leu Ala Ser Lys Thr Asn Val Tyr Ile Ser Ser Ser Ala Gly Ala Arg
            515                     520                 525
TGG CGA GAG GCA CTT CCT GGA CCT CAC TAC TAC ACA TGG GGA GAC CAC    1715
Trp Arg Glu Ala Leu Pro Gly Pro His Tyr Tyr Thr Trp Gly Asp His
530                     535                 540                 545
GGC GGA ATC ATC ACG GCC ATT GCC CAG GGC ATG GAA ACC AAC GAG CTA    1763
Gly Gly Ile Ile Thr Ala Ile Ala Gln Gly Met Glu Thr Asn Glu Leu
                550                     555                 560
AAA TAC AGT ACC AAT GAA GGG GAG ACC TGG AAA ACA TTC ATC TTC TCT    1811
Lys Tyr Ser Thr Asn Glu Gly Glu Thr Trp Lys Thr Phe Ile Phe Ser
                565                     570                 575
GAG AAG CCA GTG TTT GTG TAT GGC CTC CTC ACA GAA CCT GGG GAG AAG    1859
Glu Lys Pro Val Phe Val Tyr Gly Leu Leu Thr Glu Pro Gly Glu Lys
            580                     585                 590
AGC ACT GTC TTC ACC ATC TTT GGC TCG AAC AAA GAG AAT GTC CAC AGC    1907
Ser Thr Val Phe Thr Ile Phe Gly Ser Asn Lys Glu Asn Val His Ser
            595                     600                 605
TGG CTG ATC CTC CAG GTC AAT GCC ACG GAT GCC TTG GGA GTT CCC TGC    1955
Trp Leu Ile Leu Gln Val Asn Ala Thr Asp Ala Leu Gly Val Pro Cys
610                     615                 620                 625
ACA GAG AAT GAC TAC AAG CTG TGG TCA CCA TCT GAT GAG CGG GGG AAT    2003
Thr Glu Asn Asp Tyr Lys Leu Trp Ser Pro Ser Asp Glu Arg Gly Asn
                630                     635                 640
GAG TGT TTG CTG GGA CAC AAG ACT GTT TTC AAA CGG CGG ACC CCC CAT    2051
```

```
        Glu Cys Leu Leu Gly His Lys Thr Val Phe Lys Arg Arg Thr Pro His
                    645                 650                 655
        GCC ACA TGC TTC AAT GGA GAG GAC TTT GAC AGG CCG GTG GTC GTG TCC    2099
        Ala Thr Cys Phe Asn Gly Glu Asp Phe Asp Arg Pro Val Val Val Ser
                660                 665                 670
        AAC TGC TCC TGC ACC CGG GAG GAC TAT GAG TGT GAC TTC GGT TTC AAG    2147
        Asn Cys Ser Cys Thr Arg Glu Asp Tyr Glu Cys Asp Phe Gly Phe Lys
                675                 680                 685
        ATG AGT GAA GAT TTG TCA TTA GAG GTT TGT GTT CCA GAT CCG GAA TTT    2195
        Met Ser Glu Asp Leu Ser Leu Glu Val Cys Val Pro Asp Pro Glu Phe
        690                 695                 700                 705
        TCT GGA AAG TCA TAC TCC CCT CCT GTG CCT TGC CCT GTG GGT TCT ACT    2243
        Ser Gly Lys Ser Tyr Ser Pro Pro Val Pro Cys Pro Val Gly Ser Thr
                        710                 715                 720
        TAC AGG AGA ACG AGA GGC TAC CGG AAG ATT TCT GGG GAC ACT TGT AGC    2291
        Tyr Arg Arg Thr Arg Gly Tyr Arg Lys Ile Ser Gly Asp Thr Cys Ser
                        725                 730                 735
        GGA GGA GAT GTT GAA GCG CGA CTG GAA GGA GAG CTG GTC CCC TGT CCC    2339
        Gly Gly Asp Val Glu Ala Arg Leu Glu Gly Glu Leu Val Pro Cys Pro
                        740                 745                 750
        CTG GCA GAA GAG AAC GAG TTC ATT CTG TAT GCT GTG AGG AAA TCC ATC    2387
        Leu Ala Glu Glu Asn Glu Phe Ile Leu Tyr Ala Val Arg Lys Ser Ile
                755                 760                 765
        TAC CGC TAT GAC CTG GCC TCG GGA GCC ACC GAG CAG TTG CCT CTC ACC    2435
        Tyr Arg Tyr Asp Leu Ala Ser Gly Ala Thr Glu Gln Leu Pro Leu Thr
        770                 775                 780                 785
        GGG CTA CGG GCA GCA GTG GCC CTG GAC TTT GAC TAT GAG CAC AAC TGT    2483
        Gly Leu Arg Ala Ala Val Ala Leu Asp Phe Asp Tyr Glu His Asn Cys
                        790                 795                 800
        TTG TAT TGG TCC GAC CTG GCC TTG GAC GTC ATC CAG CGC CTC TGT TTG    2531
        Leu Tyr Trp Ser Asp Leu Ala Leu Asp Val Ile Gln Arg Leu Cys Leu
                        805                 810                 815
        AAT GGA AGC ACA GGG CAA GAG GTG ATC ATC AAT TCT GGC CTG GAG ACA    2579
        Asn Gly Ser Thr Gly Gln Glu Val Ile Ile Asn Ser Gly Leu Glu Thr
                        820                 825                 830
        GTA GAA GCT TTG GCT TTT GAA CCC CTC AGC CAG CTG CTT TAC TGG GTA    2627
        Val Glu Ala Leu Ala Phe Glu Pro Leu Ser Gln Leu Leu Tyr Trp Val
                835                 840                 845
        GAT GCA GGC TTC AAA AAG ATT GAG GTA GCT AAT CCA GAT GGC GAC TTC    2675
        Asp Ala Gly Phe Lys Lys Ile Glu Val Ala Asn Pro Asp Gly Asp Phe
        850                 855                 860                 865
        CGA CTC ACA ATC GTC AAT TCC TCT GTG CTT GAT CGT CCC AGG GCT CTG    2723
        Arg Leu Thr Ile Val Asn Ser Ser Val Leu Asp Arg Pro Arg Ala Leu
                        870                 875                 880
        GTC CTC GTG CCC CAA GAG GGG GTG ATG TTC TGG ACA GAC TGG GGA GAC    2771
        Val Leu Val Pro Gln Glu Gly Val Met Phe Trp Thr Asp Trp Gly Asp
                        885                 890                 895
        CTG AAG CCT GGG ATT TAT CGG AGC AAT ATG GAT GGT TCT GCT GCC TAT    2819
        Leu Lys Pro Gly Ile Tyr Arg Ser Asn Met Asp Gly Ser Ala Ala Tyr
                        900                 905                 910
        CAC CTG GTG TCT GAG GAT GTG AAG TGG CCC AAT GGC ATC TCT GTG GAC    2867
        His Leu Val Ser Glu Asp Val Lys Trp Pro Asn Gly Ile Ser Val Asp
                915                 920                 925
        GAC CAG TGG ATT TAC TGG ACG GAT GCC TAC CTG GAG TGC ATA GAG CGG    2915
        Asp Gln Trp Ile Tyr Trp Thr Asp Ala Tyr Leu Glu Cys Ile Glu Arg
        930                 935                 940                 945
        ATC ACG TTC AGT GGC CAG CAG CGC TCT GTC ATT CTG GAC AAC CTC CCG    2963
        Ile Thr Phe Ser Gly Gln Gln Arg Ser Val Ile Leu Asp Asn Leu Pro
                        950                 955                 960
        CAC CCC TAT GCC ATT GCT GTC TTT AAG AAT GAA ATC TAC TGG GAT GAC    3011
        His Pro Tyr Ala Ile Ala Val Phe Lys Asn Glu Ile Tyr Trp Asp Asp
                        965                 970                 975
        TGG TCA CAG CTC AGC ATA TTC CGA GCT TCC AAA TAC AGT GGG TCC CAG    3059
        Trp Ser Gln Leu Ser Ile Phe Arg Ala Ser Lys Tyr Ser Gly Ser Gln
```

99

```
                980                      985                      940
     ATG GAG ATT CTG GCA AAC CAG CTC ACG GG; CI C AT; GAC ATG AAG A'T    3107
     Met Glu Ile Leu Ala Asn Gln Leu Thr Gly Leu Met Asp Met Lys Ile
                995                      1000                     1005
     TTC TAC AAG GGG AAG AAC ACT GGA AGC AAT GCC TGT GTG CCC AGG CCA      3155
     Phe Tyr Lys Gly Lys Asn Thr Gly Ser Asn Ala Cys Val Pro Arg Pro
     1010                     1015                     1020                1025
     TGC AGC CTG CTG TGC CTG CCC AAG GCC AAC AAC AGT AGA AGC TGC AGG      3203
     Cys Ser Leu Leu Cys Leu Pro Lys Ala Asn Asn Ser Arg Ser Cys Arg
                          1030                     1035                    1040
     TGT CCA GAG GAT GTG TCC AGC AGT GTG CTT CCA TCA GGG GAC CTG ATG      3251
     Cys Pro Glu Asp Val Ser Ser Ser Val Leu Pro Ser Gly Asp Leu Met
                          1045                     1050                    1055
     TGT GAC TGC CCT CAG GGC TAT CAG CTC AAG AAC AAT ACC TGT GTC AAA      3299
     Cys Asp Cys Pro Gln Gly Tyr Gln Leu Lys Asn Asn Thr Cys Val Lys
                          1060                     1065                    1070
     GAA GAG AAC ACC TGT CTT CGC AAC CAG TAT CGC TGC AGC AAC GGG AAC      3347
     Glu Glu Asn Thr Cys Leu Arg Asn Gln Tyr Arg Cys Ser Asn Gly Asn
                     1075                     1080                    1085
     TGT ATC AAC AGC ATT TGG TGG TGT GAC TTT GAC AAC GAC TGT GGA GAC      3395
     Cys Ile Asn Ser Ile Trp Trp Cys Asp Phe Asp Asn Asp Cys Gly Asp
     1090                     1095                     1100                1105
     ATG AGC GAT GAG AGA AAC TGC CCT ACC ACC ATC TGT GAC CTG GAC ACC      3443
     Met Ser Asp Glu Arg Asn Cys Pro Thr Thr Ile Cys Asp Leu Asp Thr
                          1110                     1115                    1120
     CAG TTT CGT TGC CAG GAG TCT GGG ACT TGT ATC CCA CTG TCC TAT AAA      3491
     Gln Phe Arg Cys Gln Glu Ser Gly Thr Cys Ile Pro Leu Ser Tyr Lys
                          1125                     1130                    1135
     TGT GAC CTT GAG GAT GAC TGT GGA GAC AAC AGT GAT GAA AGT CAT TGT      3539
     Cys Asp Leu Glu Asp Asp Cys Gly Asp Asn Ser Asp Glu Ser His Cys
                     1140                     1145                    1150
     GAA ATG CAC CAG TGC CGG AGT GAC GAG TAC AAC TGC AGT TCC GGC ATG      3587
     Glu Met His Gln Cys Arg Ser Asp Glu Tyr Asn Cys Ser Ser Gly Met
                     1155                     1160                    1165
     TGC ATC CGC TCC TCC TGG GTA TGT GAC GGG GAC AAC GAC TGC AGG GAC      3635
     Cys Ile Arg Ser Ser Trp Val Cys Asp Gly Asp Asn Asp Cys Arg Asp
     1170                     1175                     1180                1185
     TGG TCT GAT GAA GCC AAC TGT ACC GCC ATC TAT CAC ACC TGT GAG GCC      3683
     Trp Ser Asp Glu Ala Asn Cys Thr Ala Ile Tyr His Thr Cys Glu Ala
                          1190                     1195                    1200
     TCC AAC TTC CAG TGC CGA AAC GGG CAC TGC ATC CCC CAG CGG TGG GCG      3731
     Ser Asn Phe Gln Cys Arg Asn Gly His Cys Ile Pro Gln Arg Trp Ala
                     1205                     1210                    1215
     TGT GAC GGG GAT ACG GAC TGC CAG GAT GGT TCC GAT GAG GAT CCA GTC      3779
     Cys Asp Gly Asp Thr Asp Cys Gln Asp Gly Ser Asp Glu Asp Pro Val
                     1220                     1225                    1230
     AAC TGT GAG AAG AAG TGC AAT GGA TTC CGC TGC CCA AAC GGC ACT TGC      3827
     Asn Cys Glu Lys Lys Cys Asn Gly Phe Arg Cys Pro Asn Gly Thr Cys
                1235                     1240                     1245
     ATC CCA TCC AGC AAA CAT TGT GAT GGT CTG CGT GAT TGC TCT GAT GGC      3875
     Ile Pro Ser Ser Lys His Cys Asp Gly Leu Arg Asp Cys Ser Asp Gly
     1250                     1255                     1260                1265
     TCC GAT GAA CAG CAC TGC GAG CCC CTC TGT ACG CAC TTC ATG GAC TTT      3923
     Ser Asp Glu Gln His Cys Glu Pro Leu Cys Thr His Phe Met Asp Phe
                          1270                     1275                    1280
     GTG TGT AAG AAC CGC CAG CAG TGC CTG TTC CAC TCC ATG GTC TGT GAC      3971
     Val Cys Lys Asn Arg Gln Gln Cys Leu Phe His Ser Met Val Cys Asp
                     1285                     1290                    1295
     GGA ATC ATC CAG TGC CGC GAC GGG TCC GAT GAG GAT GCG GCG TTT GCA      4019
     Gly Ile Ile Gln Cys Arg Asp Gly Ser Asp Glu Asp Ala Ala Phe Ala
                1300                     1305                     1310
     GGA TGC TCC CAA GAT CCT GAG TTC CAC AAG GTA TGT GAT GAG TTC GGT      4067
     Gly Cys Ser Gln Asp Pro Glu Phe His Lys Val Cys Asp Glu Phe Gly
                1315                     1320                     1325
```

```
TTC CAG TGT CAG AAT GGA GTG TGC ATC AGT TTG ATT TGG AAG TGC GAC      4115
Phe Gln Cys Gln Asn Gly Val Cys Ile Ser Leu Ile Trp Lys Cys Asp
1330                1335                1340                1345
GGG ATG GAT GAT TGC GGC GAT TAT TCT GAT GAA GCC AAC TGC GAA AAC      4163
Gly Met Asp Asp Cys Gly Asp Tyr Ser Asp Glu Ala Asn Cys Glu Asn
                1350                1355                1360
CCC ACA GAA GCC CCA AAC TGC TCC CGC TAC TTC CAG TTT CGG TGT GAG      4211
Pro Thr Glu Ala Pro Asn Cys Ser Arg Tyr Phe Gln Phe Arg Cys Glu
            1365                1370                1375
AAT GGC CAC TGC ATC CCC AAC AGA TGG AAA TGT GAC AGG GAG AAC GAC      4259
Asn Gly His Cys Ile Pro Asn Arg Trp Lys Cys Asp Arg Glu Asn Asp
            1380                1385                1390
TGT GGG GAC TGG TCT GAT GAG AAG GAT TGT GGA GAT TCA CAT ATT CTT      4307
Cys Gly Asp Trp Ser Asp Glu Lys Asp Cys Gly Asp Ser His Ile Leu
        1395                1400                1405
CCC TTC TCG ACT CCT GGG CCC TCC ACG TGT CTG CCC AAT TAC TAC CGC      4355
Pro Phe Ser Thr Pro Gly Pro Ser Thr Cys Leu Pro Asn Tyr Tyr Arg
1410                1415                1420                1425
TGC AGC AGT GGG ACC TGC GTG ATG GAC ACC TGG GTG TGC GAC GGG TAC      4403
Cys Ser Ser Gly Thr Cys Val Met Asp Thr Trp Val Cys Asp Gly Tyr
            1430                1435                1440
CGA GAT TGT GCA GAT GGC TCT GAC GAG GAA GCC TGC CCC TTG CTT GCA      4451
Arg Asp Cys Ala Asp Gly Ser Asp Glu Glu Ala Cys Pro Leu Leu Ala
                1445                1450                1455
AAC GTC ACT GCT GCC TCC ACT CCC ACC CAA CTT GGG CGA TGT GAC CGA      4499
Asn Val Thr Ala Ala Ser Thr Pro Thr Gln Leu Gly Arg Cys Asp Arg
            1460                1465                1470
TTT GAG TTC GAA TGC CAC CAA CCG AAG ACG TGT ATT CCC AAC TGG AAG      4547
Phe Glu Phe Glu Cys His Gln Pro Lys Thr Cys Ile Pro Asn Trp Lys
    1475                1480                1485
CGC TGT GAC GGC CAC CAA GAT TGC CAG GAT GGC CGG GAC GAG GCC AAT      4595
Arg Cys Asp Gly His Gln Asp Cys Gln Asp Gly Arg Asp Glu Ala Asn
1490                1495                1500                1505
TGC CCC ACA CAC AGC ACC TTG ACT TGC ATG AGC AGG GAG TTC CAG TGC      4643
Cys Pro Thr His Ser Thr Leu Thr Cys Met Ser Arg Glu Phe Gln Cys
                1510                1515                1520
GAG GAC GGG GAG GCC TGC ATT GTG CTC TCG GAG CGC TGC GAC GGC TTC      4691
Glu Asp Gly Glu Ala Cys Ile Val Leu Ser Glu Arg Cys Asp Gly Phe
                1525                1530                1535
CTG GAC TGC TCG GAC GAG AGC GAT GAA AAG GCC TGC AGT GAT GAG TTG      4739
Leu Asp Cys Ser Asp Glu Ser Asp Glu Lys Ala Cys Ser Asp Glu Leu
            1540                1545                1550
ACT GTG TAC AAA GTA CAG AAT CTT CAG TGG ACA GCT GAC TTC TCT GGG      4787
Thr Val Tyr Lys Val Gln Asn Leu Gln Trp Thr Ala Asp Phe Ser Gly
        1555                1560                1565
GAT GTG ACT TTG ACC TGG ATG AGG CCC AAA AAA ATG CCC TCT GCA TCT      4835
Asp Val Thr Leu Thr Trp Met Arg Pro Lys Lys Met Pro Ser Ala Ser
1570                1575                1580                1585
TGT GTA TAT AAT GTC TAC TAC AGG GTG GTT GGA GAG AGC ATA TGG AAG      4883
Cys Val Tyr Asn Val Tyr Tyr Arg Val Val Gly Glu Ser Ile Trp Lys
                1590                1595                1600
ACT CTG GAG ACC CAC AGC AAT AAG ACA AAC ACT GTA TTA AAA GTC TTG      4931
Thr Leu Glu Thr His Ser Asn Lys Thr Asn Thr Val Leu Lys Val Leu
            1605                1610                1615
AAA CCA GAT ACC ACG TAT CAG GTT AAA GTA CAG GTT CAG TGT CTC AGC      4979
Lys Pro Asp Thr Thr Tyr Gln Val Lys Val Gln Val Gln Cys Leu Ser
            1620                1625                1630
AAG GCA CAC AAC ACC AAT GAC TTT GTG ACC CTG AGG ACC CCA GAG GGA      5027
Lys Ala His Asn Thr Asn Asp Phe Val Thr Leu Arg Thr Pro Glu Gly
        1635                1640                1645
TTG CCA GAT GCC CCT CGA AAT CTC CAG CTG TCA CTC CCC AGG GAA GCA      5075
Leu Pro Asp Ala Pro Arg Asn Leu Gln Leu Ser Leu Pro Arg Glu Ala
1650                1655                1660                1665
GAA GGT GTG ATT GTA GGC CAC TGG GCT CCT CCC ATC CAC ACC CAT GGC      5123
Glu Gly Val Ile Val Gly His Trp Ala Pro Pro Ile His Thr His Gly
```

```
       Glu Gly Val Ile Val Gly His Trp Ala Pro Pro Ile His Thr His Gly
                     1670                  1675                  1680
       CTC ATC CGT GAG TAC ATT GTA GAA TAC AGC AGG AGT GGT TCC AAG ATG    5171
       Leu Ile Arg Glu Tyr Ile Val Glu Tyr Ser Arg Ser Gly Ser Lys Met
                     1685                  1690                  1695
       TGG GCC TCC CAG AGG GCT GCT AGT AAC TTT ACA GAA ATC AAG AAC TTA    5219
       Trp Ala Ser Gln Arg Ala Ala Ser Asn Phe Thr Glu Ile Lys Asn Leu
                     1700                  1705                  1710
       TTG GTC AAC ACT CTA TAC ACC GTC AGA GTG GCT GCG GTG ACT AGT CGT    5267
       Leu Val Asn Thr Leu Tyr Thr Val Arg Val Ala Ala Val Thr Ser Arg
                     1715                  1720                  1725
       GGA ATA GGA AAC TGG AGC GAT TCT AAA TCC ATT ACC ACC ATA AAA GGA    5315
       Gly Ile Gly Asn Trp Ser Asp Ser Lys Ser Ile Thr Thr Ile Lys Gly
       1730                  1735                  1740                  1745
       AAA GTG ATC CCA CCA CCA GAT ATC CAC ATT GAC AGC TAT GGT GAA AAT    5363
       Lys Val Ile Pro Pro Pro Asp Ile His Ile Asp Ser Tyr Gly Glu Asn
                     1750                  1755                  1760
       TAT CTA AGC TTC ACC CTG ACC ATG GAG AGT GAT ATC AAG GTG AAT GGC    5411
       Tyr Leu Ser Phe Thr Leu Thr Met Glu Ser Asp Ile Lys Val Asn Gly
                     1765                  1770                  1775
       TAT GTG GTG AAC CTT TTC TGG GCA TTT GAC ACC CAC AAG CAA GAG AGG    5459
       Tyr Val Val Asn Leu Phe Trp Ala Phe Asp Thr His Lys Gln Glu Arg
                     1780                  1785                  1790
       AGA ACT TTG AAC TTC CGA GGA AGC ATA TTG TCA CAC AAA GTT GGC AAT    5507
       Arg Thr Leu Asn Phe Arg Gly Ser Ile Leu Ser His Lys Val Gly Asn
                     1795                  1800                  1805
       CTG ACA GCT CAT ACA TCC TAT GAG ATT TCT GCC TGG GCC AAG ACT GAC    5555
       Leu Thr Ala His Thr Ser Tyr Glu Ile Ser Ala Trp Ala Lys Thr Asp
       1810                  1815                  1820                  1825
       TTG GGG GAT AGC CCT CTG GCA TTT GAG CAT GTT ATG ACC AGA GGG GTT    5603
       Leu Gly Asp Ser Pro Leu Ala Phe Glu His Val Met Thr Arg Gly Val
                     1830                  1835                  1840
       CGC CCA CCT GCA CCT AGC CTC AAG GCC AAA GCC ATC AAC CAG ACT GCA    5651
       Arg Pro Pro Ala Pro Ser Leu Lys Ala Lys Ala Ile Asn Gln Thr Ala
                     1845                  1850                  1855
       GTG GAA TGT ACC TGG ACC GGC CCC CGG AAT GTG GTT TAT GGT ATT TTC    5699
       Val Glu Cys Thr Trp Thr Gly Pro Arg Asn Val Val Tyr Gly Ile Phe
                     1860                  1865                  1870
       TAT GCC ACG TCC TTT CTT GAC CTC TAT CGC AAC CCG AAG AGC TTG ACT    5747
       Tyr Ala Thr Ser Phe Leu Asp Leu Tyr Arg Asn Pro Lys Ser Leu Thr
                     1875                  1880                  1885
       ACT TCA CTC CAC AAC AAG ACG GTC ATT GTC AGT AAG GAT GAG CAG TAT    5795
       Thr Ser Leu His Asn Lys Thr Val Ile Val Ser Lys Asp Glu Gln Tyr
       1890                  1895                  1900                  1905
       TTG TTT CTG GTC CGT GTA GTG GTA CCC TAC CAG GGG CCA TCC TCT GAC    5843
       Leu Phe Leu Val Arg Val Val Val Pro Tyr Gln Gly Pro Ser Ser Asp
                     1910                  1915                  1920
       TAC GTT GTA GTG AAG ATG ATC CCG GAC AGC AGG CTT CCA CCC CGT CAC    5891
       Tyr Val Val Val Lys Met Ile Pro Asp Ser Arg Leu Pro Pro Arg His
                     1925                  1930                  1935
       CTG CAT GTG GTT CAT ACG GGC AAA ACC TCC GTG GTC ATC AAG TGG GAA    5939
       Leu His Val Val His Thr Gly Lys Thr Ser Val Val Ile Lys Trp Glu
       1940                  1945                  1950
       TCA CCG TAT GAC TCT CCT GAC CAG GAC TTG TTG TAT GCA ATT GCA GTC    5987
       Ser Pro Tyr Asp Ser Pro Asp Gln Asp Leu Leu Tyr Ala Ile Ala Val
                     1955                  1960                  1965
       AAA GAT CTC ATA AGA AAG ACT GAC AGG AGC TAC AAA GTA AAA TCC CGT    6035
       Lys Asp Leu Ile Arg Lys Thr Asp Arg Ser Tyr Lys Val Lys Ser Arg
       1970                  1975                  1980                  1985
       AAC AGC ACT GTG GAA TAC ACC CTT AAC AAG TTG GAG CCT GGC GGG AAA    6083
       Asn Ser Thr Val Glu Tyr Thr Leu Asn Lys Leu Glu Pro Gly Gly Lys
                     1990                  1995                  2000
       TAC CAC ATC ATT GTC CAA CTG GGG AAC ATG AGC AAA GAT TCC AGC ATA    6131
       Tyr His Ile Ile Val Gln Leu Gly Asn Met Ser Lys Asp Ser Ser Ile
```

```
                    2005                    2010                    2015
AAA ATT ACC ACA GTT TCA TTA TCA GCA CCT GAT GCC TTA AAA ATC ATA   6179
Lys Ile Thr Thr Val Ser Leu Ser Ala Pro Asp Ala Leu Lys Ile Ile
                    2020                    2025                    2030
ACA GAA AAT GAT CAT GTT CTT CTG TTT TGG AAA AGC CTG GCT TTA AAG   6227
Thr Glu Asn Asp His Val Leu Leu Phe Trp Lys Ser Leu Ala Leu Lys
                    2035                    2040                    2045
GAA AAG CAT TTT AAT GAA AGC AGG GGC TAT GAG ATA CAC ATG TTT GAT   6275
Glu Lys His Phe Asn Glu Ser Arg Gly Tyr Glu Ile His Met Phe Asp
     2050                    2055                    2060          2065
AGT GCC ATG AAT ATC ACA GCT TAC CTT GGG AAT ACT ACT GAC AAT TTC   6323
Ser Ala Met Asn Ile Thr Ala Tyr Leu Gly Asn Thr Thr Asp Asn Phe
                    2070                    2075          2080
TTT AAA ATT TCC AAC CTG AAG ATG GGT CAT AAT TAC ACG TTC ACC GTC   6371
Phe Lys Ile Ser Asn Leu Lys Met Gly His Asn Tyr Thr Phe Thr Val
                    2085                    2090          2095
CAA GCA AGA TGC CTT TTT GGC AAC CAG ATC TGT GGG GAG CCT GCC ATC   6419
Gln Ala Arg Cys Leu Phe Gly Asn Gln Ile Cys Gly Glu Pro Ala Ile
          2100                    2105                    2110
CTG CTG TAC GAT GAG CTG GGG TCT GGT GCA GAT GCA TCT GCA ACG CAG   6467
Leu Leu Tyr Asp Glu Leu Gly Ser Gly Ala Asp Ala Ser Ala Thr Gln
          2115                    2120                    2125
GCT GCC AGA TCT ACG GAT GTT GCT GCT GTG GTG GTG CCC ATC TTA TTC   6515
Ala Ala Arg Ser Thr Asp Val Ala Ala Val Val Val Pro Ile Leu Phe
     2130                    2135                    2140          2145
CTG ATA CTG CTG AGC CTG GGG GTG GGG TTT GCC ATC CTG TAC ACG AAG   6563
Leu Ile Leu Leu Ser Leu Gly Val Gly Phe Ala Ile Leu Tyr Thr Lys
                    2150                    2155                    2160
CAC CGG AGG CTG CAG AGC AGC TTC ACC GCC TTC GCC AAC AGC CAC TAC   6611
His Arg Arg Leu Gln Ser Ser Phe Thr Ala Phe Ala Asn Ser His Tyr
                    2165                    2170                    2175
AGC TCC AGG CTG GGG TCC GCA ATC TTC TCC TCT GGG GAT GAC CTG GGG   6659
Ser Ser Arg Leu Gly Ser Ala Ile Phe Ser Ser Gly Asp Asp Leu Gly
          2180                    2185                    2190
GAA GAT GAT GAA GAT GCC CCT ATG ATA ACT GGA TTT TCA GAT GAC GTC   6707
Glu Asp Asp Glu Asp Ala Pro Met Ile Thr Gly Phe Ser Asp Asp Val
          2195                    2200                    2205
CCC ATG GTG ATA GCC TGAAAGAGCT TTCCTCACTA GAAACCAAAT GGTGTAAATA   6762
Pro Met Val Ile Ala
2210
TTTTATTTGA TAAAGATAGT TGATGGTTTA TTTTAAAAGA TGCACTTTGA GTTGCAATAT 6822
GTTATTTTTA TATGGGCCAA A                                          6843
```

Claims

1. DNA having a nucleotide sequence as shown by Sequence ID No. 1.

2. An LDL receptor analog protein having an amino acid sequence as shown by Sequence ID No. 2 and coded by the DNA of Claim 1.

3. DNA having a nucleotide sequence as shown by Sequence ID No. 5.

4. An LDL receptor analog protein having an amino acid sequence as shown by Sequence ID No. 6 and coded by the DNA of Claim 3.

5. A recombinant vector comprising DNA as shown by Sequence ID No. 1 or 5 and a replicable vector.

6. Transformant cells which harbor the recombinant vector of Claim 5.

7. A method for the production of an LDL receptor analog protein comprising the steps of culturing the transformants of Claim 6 and collecting a polypeptide produced in the culture.

5

10

15

20

25

30

35

40

45

50

55

(72) Inventors:
• Saito, Yasushi
Chiba-shi, Chiba (JP)
• Iwasaki, Akio
Tsuchiura-shi, Ibaraki (JP)
• Arai, Koichi
Urawa-shi, Saitama (JP)
• Yamazaki, Hiroyuki
Higashimurayama-shi, Tokyo (JP)

(74) Representative:
Wächtershäuser, Günter, Prof. Dr.
Patentanwalt,
Tal 29
80331 München (DE)

(54) **Novel LDL receptor analog protein and the gene coding therefor**

(57) The present invention is drawn to the gene of a novel LDL receptor family receptor which participates in lipoprotein metabolism, a critical factor that triggers the onset of arteriosclerosis.

The invention provides DNA having a nucleotide sequence as shown by Sequence ID No. 1 or No.5 is disclosed as well as rabbit tissue or human tissue LDL receptor analog protein having an amino acid sequence of Sequence ID No. 2 or 6 coded by such DNA.

EP 0 773 290 A3

European Patent Office

EUROPEAN SEARCH REPORT

Application Number

EP 96 11 6108

DOCUMENTS CONSIDERED TO BE RELEVANT

Category	Citation of document with indication, where appropriate, of relevant passages	Relevant to claim	CLASSIFICATION OF THE APPLICATION (Int.Cl.6)
A	WO 94 10322 A (UNIVERSITY OF TEXAS) * the whole document * ---	1-7	C12N15/12 C07K14/705 //C12N15/70, C12N15/79
A	WO 95 13374 A (BAYLOR COLLEGE OF MEDICINE) * the whole document * ---	1-7	
A	KOZARSKY K. ET AL.: "In vivo correction of low density lipoprotein receptor deficiency in the Watanabe heritable hyperlipemic rabbit with recombinant adenoviruses" JOURNAL OF BIOLOGICAL CHEMISTRY, vol. 269, no. 18, 6 May 1994, MD US, pages 13695-13702, XP002062181 * the whole document * ---	1-7	
A	WILSON J.M. ET AL.: "Hepathocyte-directed gene transfer in vivo leads to transient improvement of hypercholesterolemia in low density lipoprotein receptor-deficient rabbits" JOURNAL OF BIOLOGICAL CHEMISTRY, vol. 267, no. 2, 15 January 1995, MD US, pages 963-967, XP002062183 * the whole document * -----	1-7	**TECHNICAL FIELDS SEARCHED (Int.Cl.6)** C07K

The present search report has been drawn up for all claims

Place of search	Date of completion of the search	Examiner
BERLIN	16 April 1998	Panzica, G

EPO FORM 1503 03 82 (P04C01)